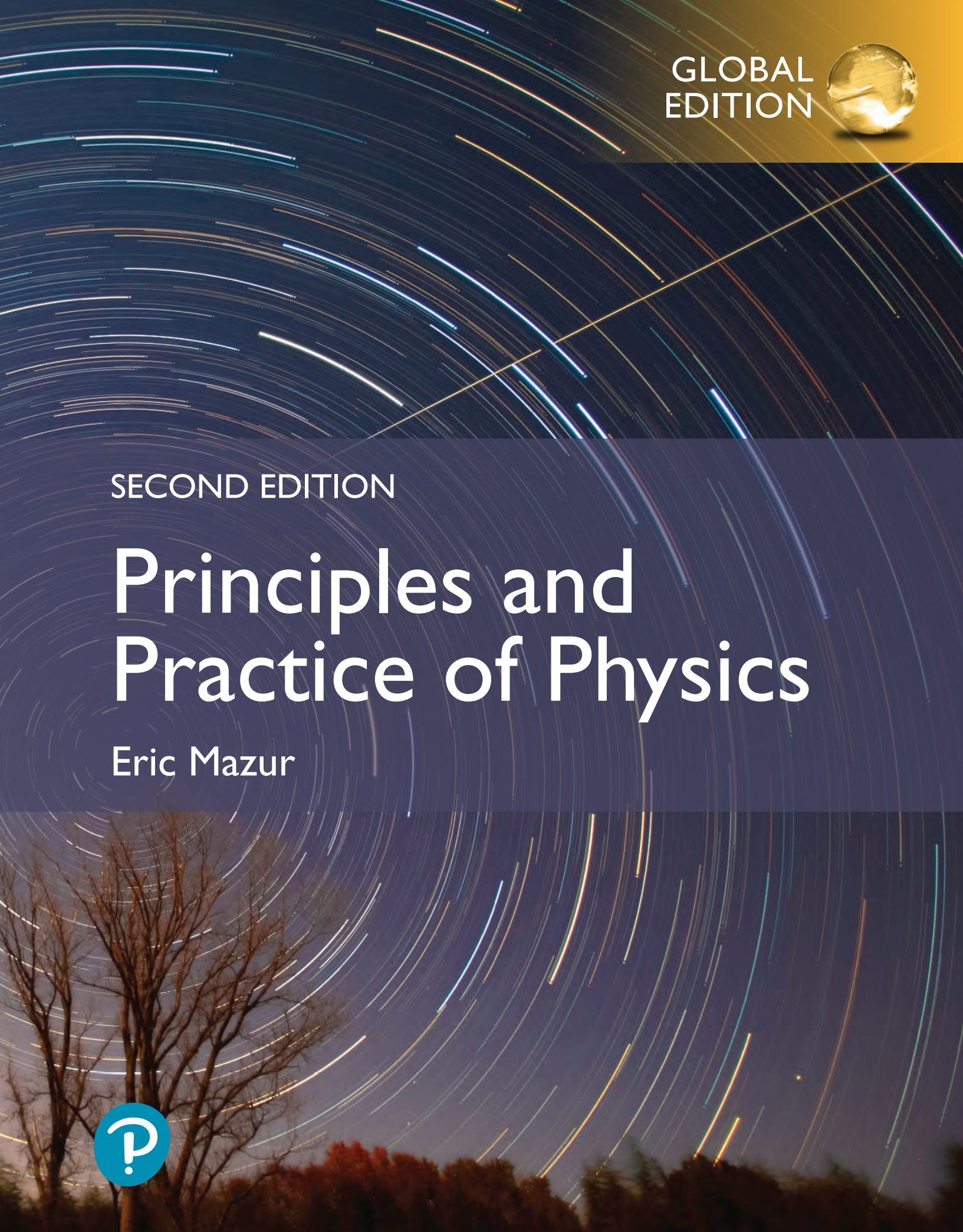

GLOBAL
EDITION
SECOND EDITION
Principles and
Practice of Physics
Eric Mazur

SECOND EDITION

GLOBAL EDITION

PRINCIPLES & PRACTICE OF PHYSICS

ERIC MAZUR

HARVARD UNIVERSITY

Pearson

Pearson Education Limited
KAO Two
KAO Park
Hockham Way
Harlow
CM17 9SR
United Kingdom

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

© Pearson Education Limited 2022

The rights of Eric Mazur to be identified as the author of this work have been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

Authorized adaptation from the United States edition, entitled Principles and Practice of Physics*, 2nd Edition, ISBN 978-0-13561086-2 by Eric Mazur, published by Pearson Education © 2021.*

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a license permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS. For information regarding permissions, request forms and the appropriate contacts within the Pearson Education Global Rights & Permissions department, please visit www.pearsoned.com/permissions/.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

Acknowledgements of third-party content appear on page 1335, which constitutes an extension of this copyright page.

PEARSON, ALWAYS LEARNING, and Mastering™ Chemistry are exclusive trademarks in the U.S. and/or other countries owned by Pearson Education, Inc. or its affiliates.

Unless otherwise indicated herein, any third-party trademarks that may appear in this work are the property of their respective owners and any references to third-party trademarks, logos or other trade dress are for demonstrative or descriptive purposes only. Such references are not intended to imply any sponsorship, endorsement, authorization, or promotion of Pearson's products by the owners of such marks, or any relationship between the owner and Pearson Education, Inc. or its affiliates, authors, licensees or distributors.

ISBN 10: 1-292-36470-X
ISBN 13: 978-1-292-36470-4
eBook ISBN 13: 978-1-292-36472-8

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

1 22

Typeset by Straive

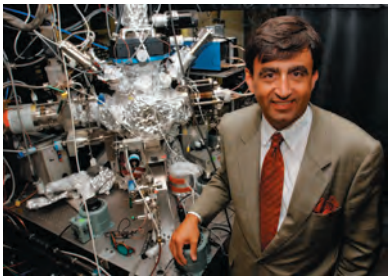
eBook formatted by B2R Technologies Pvt. Ltd.

Brief Contents

Volume 1 of *Principles & Practice of Physics* includes Chapters 1–21. Volume 2 of *Principles & Practice of Physics* includes Chapters 22–34.

CHAPTER 1 Foundations 19
CHAPTER 2 Motion in One Dimension 53
CHAPTER 3 Acceleration 86
CHAPTER 4 Momentum 116
CHAPTER 5 Energy 149
CHAPTER 6 Principle of Relativity 176
CHAPTER 7 Interactions 211
CHAPTER 8 Force 246
CHAPTER 9 Work 282
CHAPTER 10 Motion in a Plane 314
CHAPTER 11 Motion in a Circle 351
CHAPTER 12 Torque 388
CHAPTER 13 Gravity 425
CHAPTER 14 Special Relativity 461
CHAPTER 15 Periodic Motion 504
CHAPTER 16 Waves in One Dimension 539
CHAPTER 17 Waves in Two and Three Dimensions 580
CHAPTER 18 Fluids 616
CHAPTER 19 Entropy 664
CHAPTER 20 Energy Transferred Thermally 701
CHAPTER 21 Degradation of Energy 742
CHAPTER 22 Electric Interactions 783
CHAPTER 23 The Electric Field 813
CHAPTER 24 Gauss's Law 845
CHAPTER 25 Work and Energy in Electrostatics 878
CHAPTER 26 Charge Separation and Storage 908
CHAPTER 27 Magnetic Interactions 941
CHAPTER 28 Magnetic Fields of Charged Particles in Motion 976
CHAPTER 29 Changing Magnetic Fields 1011
CHAPTER 30 Changing Electric Fields 1042
CHAPTER 31 Electric Circuits 1079
CHAPTER 32 Electronics 1121
CHAPTER 33 Ray Optics 1164
CHAPTER 34 Wave and Particle Optics 1207

About the Author



Eric Mazur is the Balkanski Professor of Physics and Applied Physics at Harvard University, Area Chair of Applied Physics, and a member of the faculty of Education at the Harvard Graduate School of Education. Dr. Mazur is a renowned scientist and researcher in optical physics and in education research, and a sought-after author and speaker.

Dr. Mazur joined the faculty at Harvard shortly after obtaining his Ph.D. at the University of Leiden in the Netherlands. He was awarded Honorary Doctorates from the École Polytechnique and the University of Montreal, the Universidad Nacional Mayor de San Marcos in Lima, Peru, and the Katholieke Universiteit Leuven in Belgium. Dr. Mazur holds honorary professorships at the Institute of Semiconductor Physics of the Chinese Academy of Sciences in Beijing, the Institute of Laser Engineering at the Beijing University of Technology, the Beijing Normal University, Sichuan University, and Nanjing University of Science and Technology. He is a Member of the Royal Academy of Sciences of the Netherlands and a Member of the Royal Holland Society of Sciences and Humanities. In 2014, Dr. Mazur became the inaugural recipient of the Minerva Prize, and in 2018 he received the inaugural International Flipped Learning Award from the American Academy of Learning Arts and Sciences.

Dr. Mazur has held appointments as Visiting Professor or Distinguished Lecturer at Carnegie Mellon University, the Ohio State University, the Pennsylvania State University, Princeton University, Vanderbilt University, Hong Kong University, the University of Leuven in Belgium, and National Taiwan University in Taiwan, among others. From 2015–2017 Dr. Mazur served as Vice-President, President-Elect, and President of the Optical Society.

In addition to his work in optical physics, Dr. Mazur is interested in education, science policy, outreach, and the public perception of science. In 1990, he began developing Peer instruction, a method for teaching large lecture classes interactively. This teaching method has developed a large following, both nationally and internationally, and has been adopted across many science disciplines.

Dr. Mazur is author or co-author of over 300 scientific publications and holds three dozen patents. He has also written on education and is the author of *Peer Instruction: A User's Manual* (Pearson, 1997), a book that explains how to teach large lecture classes interactively. In 2006, he helped produce the award-winning DVD *Interactive Teaching*. He is the co-founder of Learning Catalytics, a platform for promoting interactive problem solving in the classroom, and of Perusall, the first truly AI-driven social learning platform.

To the Student

Let me tell you a bit about myself.

I always knew exactly what I wanted to do. It just never worked out that way.

When I was seven years old, my grandfather gave me a book about astronomy. Growing up in the Netherlands I became fascinated by the structure of the solar system, the Milky Way, the universe. I remember struggling with the concept of infinite space and asking endless questions without getting satisfactory answers. I developed an early passion for space and space exploration. I knew I was going to be an astronomer. In high school I was good at physics, but when I entered university and had to choose a major, I chose astronomy.

It took only a few months for my romance with the heavens to unravel. Instead of teaching me about the mysteries and structure of the universe, astronomy had been reduced to a mind-numbing web of facts, from declinations and right ascensions to semi-major axes and eccentricities. Disillusioned about astronomy, I switched majors to physics. Physics initially turned out to be no better than astronomy, and I struggled to remain engaged. I managed to make it through my courses, often by rote memorization, but the beauty of science eluded me.

It wasn't until doing research in graduate school that I rediscovered the beauty of science. I knew one thing for sure, though: I was never going to be an academic. I was going to do something useful in my life. Just before obtaining my doctorate, I lined up my dream job working on the development of the compact disc, but I decided to spend one year doing postdoctoral research first.

It was a long year. After my postdoc, I accepted a junior faculty position and started teaching. That's when I discovered that the combination of doing research—uncovering the mysteries of the universe—and teaching—helping others to see the beauty of the universe—is a wonderful combination.

When I started teaching, I did what all teachers did at the time: lecture. It took almost a decade to discover that my award-winning lecturing did for my students exactly what the courses I took in college had done for me: It turned the subject that I was teaching into a collection of facts that my students memorized by rote. Instead of transmitting the beauty of my field, I was essentially regurgitating facts to my students.

When I discovered that my students were not mastering even the most basic principles, I decided to completely change my approach to teaching. Instead of lecturing, I asked students to read my lecture notes at home, and then, in class, I taught by questioning—by asking my students to reflect on concepts, discuss in pairs, and experience their own "aha!" moments.

Over the course of more than twenty years, the lecture notes have evolved into this book. Consider this book to be my best possible "lecturing" to you. But instead of listening to me without having the opportunity to reflect and think, this book will permit you to pause and think; to hopefully experience many "aha!" moments on your own.

I hope this book will help you develop the thinking skills that will make you successful in your career. And remember: your future may be—and likely will be—very different from what you imagine.

I welcome any feedback you have. Feel free to send me email or tweets.

I wrote this book for you.

Eric Mazur
@eric_mazur
mazur@harvard.edu
Cambridge, MA

To the Instructor

As you may recall from the first edition of this book, the idea of using conservation principles derived from a conversation with a dear friend and colleague, Albert Altman, professor at the University of Massachusetts, Lowell, who asked me if I was familiar with the approach to physics taken by Ernst Mach.

Mach treats conservation of momentum before discussing the laws of motion. It involves direct experimental observation, which appealed to me. His formulation of mechanics had a profound influence on Einstein. Most physicists never use the concept of force because it relates only to mechanics. It has no role in quantum physics, for example. The conservation principles, however, hold throughout all of physics. In that sense they are much more fundamental than Newton's laws. Furthermore, conservation principles involve only algebra, whereas Newton's second law is a differential equation.

Physics education research has shown that the concept of force, where most physics books begin, is fraught with pitfalls. What's more, after tediously deriving many results using kinematics and dynamics, most physics textbooks show that you can derive the same results from conservation principles in just one or two lines. Why not do the easy way first?

In this edition Principles and Practice of Physics, I start with conservation of both momentum and energy, and later bring in the concept of force. The approach is more unified and modern—the conservation principles are the theme that runs throughout this entire book.

Additional motives for writing this text came from my own teaching. Most textbooks focus on the acquisition of information and on the development of procedural knowledge. This focus comes at the expense of conceptual understanding or the ability to transfer knowledge to a new context. As explained below, I have structured this text to redress that balance. I also have drawn deeply on the results of physics education research, including that of my own research group.

Organization of this book

As I considered the best way to convey the conceptual framework of mechanics, it became clear that the standard curriculum needed rethinking. For example, standard texts are forced to redefine certain concepts more than once—a strategy that we know befuddles students. (Examples are work, the standard definition of which is incompatible with the first law of thermodynamics, and energy, which is redefined when modern physics is discussed.)

Another point that has always bothered me is the arbitrary division between "modern" and "classical" physics. In most texts, the first thirty-odd chapters present physics essentially as it was known at the end of the 19th century; "modern physics" gets tacked on at the end. There's no need for this separation. Our goal should be to explain physics in the way that works best for students, using our full contemporary understanding. All physics is modern!

That is why my table of contents departs from the "standard organization" in the following specific ways.

Emphasis on conservation laws. As mentioned earlier, this book introduces the conservation laws early and treats them the way they should be: as the backbone of physics. The advantages of this shift are many. First, it avoids many of the standard pitfalls related to the concept of force, and it leads naturally to the two-body character of forces and the laws of motion. Second, the conservation laws enable students to solve a wide variety of problems without any calculus. Indeed, for complex systems, the conservation laws are often the natural (or only) way to solve problems. Third, the book deduces the conservation laws from experimental observations, helping to make clear their connection with the world around us. I and several other instructors have tested this approach extensively in our classes and found markedly improved performance on problems involving momentum and energy, with large gains on assessment instruments like the Force Concept Inventory.

Early emphasis on the concept of system. Fundamental to most physical models is the separation of a system from its environment. This separation is so basic that physicists tend to carry it out unconsciously, and traditional texts largely gloss over it. This text introduces the concept in the context of conservation principles and uses it consistently.

Postponement of vectors. Most introductory physics concerns phenomena that take place along one dimension. Problems that involve more than one dimension can be broken down into one-dimensional problems using vectorial notation. So, a solid understanding of physics in one dimension is of fundamental importance. However, by introducing vectors in more than one dimension from the start, standard texts distract the student from the basic concepts of kinematics.

In this book, I develop the complete framework of mechanics for motions and interactions in one dimension. I introduce the second dimension when it is needed, starting with rotational motion. Hence, students can focus on the actual physics early on.

Table 1 Scheduling matrix

Topic	Chapters	Can be inserted after chapter . . .	Chapters that can be omitted without affecting continuity
Mechanics	1–14		6, 13–14
Waves	15–17	12	16–17
Fluids	18	9	
Thermal Physics	19–21	10	21
Electricity & Magnetism	22–30	12 (but 17 is needed for 29–30)	29–30
Circuits	31–32	26 (but 30 is needed for 32)	32
Optics	33–34	17	34

Just-in-time introduction of concepts. Wherever possible, I introduce concepts only when they are necessary. This approach allows students to put ideas into immediate practice, leading to better assimilation.

Integration of modern physics. A survey of syllabi shows that less than half the calculus-based courses in the United States cover modern physics. I have therefore integrated selected "modern" topics throughout the text. For example, special relativity is covered in Chapter 14, at the end of mechanics. Chapter 32, Electronics, includes sections on semiconductors and semiconductor devices. Chapter 34, Wave and Particle Optics, contains sections on quantization and photons.

Modularity. I have written the book in a modular fashion so it can accommodate a variety of curricula (See Table 1, "Scheduling matrix").

The book contains two major parts, Mechanics and Electricity and Magnetism, plus five shorter parts. The two major parts by themselves can support an in-depth two-semester or three-quarter course that presents a complete picture of physics embodying the fundamental ideas of modern physics. Additional parts can be added for a longer or faster-paced course. The five shorter parts are more or less self-contained, although they do build on previous material, so their placement is flexible. Within each part or chapter, more advanced or difficult material is placed at the end.

Pedagogy

This text draws on many models and techniques derived from my own teaching and from physics education research. The following are major themes that I have incorporated throughout.

Separation of conceptual and mathematical frameworks. Each chapter is divided into two parts: Concepts and Quantitative Tools. The first part, Concepts, develops the full conceptual framework of the topic and addresses many of the common questions students have. It concentrates on the underlying ideas and paints the big picture, whenever possible without equations. The second part of the chapter, Quantitative Tools, then develops the mathematical framework.

Deductive approach; focus on ideas before names and equations. To the extent possible, this text develops arguments deductively, starting from observations, rather than stating principles and then "deriving" them. This approach makes the material easier to assimilate for students. In the same vein, this text introduces and explains each idea before giving it a formal name or mathematical definition.

Stronger connection to experiment and experience. Physics stems from observations, and this text is structured so that it can do the same. As much as possible, I develop the material from experimental observations (and preferably those that students can make) rather than assertions. Most chapters use actual data in developing ideas, and new notions are always introduced by going from the specific to the general—whenever possible by interpreting everyday examples.

By contrast, standard texts often introduce laws in their most general form and then show that these laws are consistent with specific (and often highly idealized) cases. Consequently, the world of physics and the "real" world remain two different things in the minds of students.

Addressing physical complications. I also strongly oppose presenting unnatural situations; real life complications must always be confronted head-on. For example, the use of unphysical words like frictionless or massless sends a message to the students that physics is unrealistic or, worse, that the world of physics and the real world are unrelated entities. This can easily be avoided by pointing out that friction or mass may be neglected under certain circumstances and pointing out why this may be done.

Engaging the student. Education is more than just transfer of information. Engaging the student's mind so the information can be assimilated is essential. To this end, the text is written as a dialog between author and reader (often invoking the reader—you—in examples)

and is punctuated by Checkpoints—questions that require the reader to stop and think. The text following a Checkpoint often refers directly to its conclusions. Students will find complete solutions to all the Checkpoints at the back of the book; these solutions are written to emphasize physical reasoning and discovery.

Visualization. Visual representations are central to physics, so I developed each chapter by designing the figures before writing the text. Many figures use multiple representations to help students make connections (for example, a sketch may be combined with a graph and a bar diagram). Also, in accordance with research, the illustration style is spare and simple, putting the emphasis on the ideas and relationships rather than on irrelevant details. The figures do not use perspective unless it is needed, for instance.

Physics for today's student

This new edition focuses on today's physics student who not only learns in the physical classroom but gleans knowledge in a digital environment. The content format is modified for students to actively engage with online content first. The second edition optimizes the delivery of the content by combining both volumes of the first edition (Principles volume and Practice volume) and providing it as one single volume via etext and Mastering Physics in addition to the print book.

Best practices reflect the idea that by engaging with the material before coming to class better prepares students to learn. The new edition provides new pre-lecture videos by both Eric Mazur and his Harvard colleague Greg Kestin (researcher and consultant for NOVA) to help students come to class ready to participate, encourage the understanding of real-world application of physics, and support instructors in building active and relevant classes.

As pointed out earlier, each chapter is divided into two parts. The first part (Concepts) develops the conceptual framework in an accessible way, relying primarily on qualitative descriptions and illustrations. In addition to including Checkpoints, each Concepts section ends with a one-page Self-quiz consisting of qualitative questions. The second part of each chapter (Quantitative Tools) formalizes the ideas developed in the first part in mathematical terms. While concise, it is relatively traditional in nature—teachers should be able to continue to use material developed for earlier courses. To avoid creating the impression that equations are more important than the concepts behind them, no equations are highlighted or boxed. Both parts of the chapters contain worked examples to help students develop problem-solving skills.

At the end of each chapter is a Chapter Summary and a Questions and Problems section. The problems 1) offer a range of levels; 2) include problems relating to client disciplines (life sciences, engineering, chemistry, astronomy, etc.); 3) use the second person as much as possible to draw in the student; and 4) do not spoon-feed the students with information and unnecessary diagrams. The problems are classified into three levels as follows: (•) application of single concept; numerical plug-and-chug; (••) nonobvious application of single concept or application of multiple concepts from current chapter; straightforward numerical or algebraic computation; (•••) application of multiple concepts, possibly spanning multiple chapters. Context-rich problems are designated CR.

Additional material can be found online in Mastering Physics:

1. *Review Questions*. The goal of this section is to allow students to quickly review the corresponding chapter. The questions are straightforward one-liners starting with "what" and "how" (rather than "why" or "what if"). These questions are in Mastering Physics.
2. *Developing a Feel*. The goals of this section are to develop a quantitative feel for the quantities introduced in the chapter; to connect the subject of the chapter to the real world; to train students in making estimates and assumptions; to bolster students' confidence in dealing with unfamiliar material. It can be used for self-study or for a homework or recitation assignment. This section, which has no equivalent in existing books, combines a number of ideas (specifically, Fermi problems and tutoring in the style of the *Princeton Learning Guide*). The idea is to start with simple estimation problems and then build up to Fermi problems (in early chapters Fermi problems are hard to compose because few concepts have been introduced). Because students initially find these questions hard, the section provides many hints, which take the form of guiding questions. A key then provides answers to these "hints." These Developing a Feel questions are now included in Mastering Physics.
3. *Worked and Guided Problems*. This section contains worked examples whose primary goal is to teach problem solving. The Worked Problems are fully solved; the Guided Problems have a list of guiding questions and suggestions to help the student think about how to solve the problem. Typically, each Worked Problem is followed by a related Guided Problem. Both are available in Mastering Physics.

Instructor supplements

Downloadable ***Instructor Resources*** includes an Image Library, **Lecture Outlines** with embedded **Clicker Questions in PowerPoint®**, and the ***Instructor's Solutions Manual***.

The ***Instructor's Solutions Manual*** is a comprehensive solutions manual containing complete answers and solutions to Questions and Problems found at the end of each chapter, as well as all Developing a Feel questions and Guided Problems found in Mastering Physics. The solutions to the Guided Problems use the book's four-step problem-solving strategy (Getting Started, Devise Plan, Execute Plan, Evaluate Result).

Mastering Physics® is the leading online homework, tutorial, and assessment product designed to improve results by helping students quickly master concepts. Students benefit from self-paced tutorials that feature specific wrong-answer feedback, hints, and a wide variety of educationally effective content to keep them engaged and on track. Robust diagnostics and unrivalled gradebook reporting allow instructors to pinpoint the weaknesses and misconceptions of a student or class to provide timely intervention.

Mastering Physics enables instructors to:

- Easily assign **tutorials** that provide individualized coaching.
- Mastering's hallmark **Hints** and **Feedback** offer scaffolded instruction similar to what students would experience in an office hour.
- **Hints** (declarative and Socratic) can provide problem-solving strategies or break the main problem into simpler exercises.
- **Feedback** lets the student know precisely what misconception or misunderstanding is evident from their answer and offers ideas to consider when attempting the problem again.

Learning Catalytics™ is a "bring your own device" student engagement, assessment, and classroom intelligence system available within Mastering Physics. With Learning Catalytics you can:

- Assess students in real time, using open-ended tasks to probe student understanding.
- Understand immediately where students are and adjust your lecture accordingly.
- Improve your students' critical-thinking skills.
- Access rich analytics to understand student performance.
- Add your own questions to make Learning Catalytics fit your course exactly.
- Manage student interactions with intelligent grouping and timing.

The **Test Bank** contains more than 2000 high-quality problems, with a range of multiple-choice, true-false, short-answer, and conceptual questions correlated to *Principles & Practice of Physics* chapters. Test files are provided in both TestGen® and Microsoft® Word for Mac and PC.

Instructor supplements are available in the Instructor Resource area of Mastering Physics (www.masteringphysics.com).

Student supplements

Mastering Physics (www.masteringphysics.com) is designed to provide students with customized coaching and individualized feedback to help improve problem-solving skills. Students complete homework efficiently and effectively with tutorials that provide targeted help. By combining trusted author content with digital tools developed to engage students and emulate the office-hour experience, Mastering personalizes learning and improves results for each student. Built for, and directly tied to the text, Mastering Physics gives students a platform to practice, learn, and apply knowledge outside of the classroom.

Acknowledgments

This book would not exist without the contributions from many people. It was Tim Bozik, currently the President of Global Product at Pearson, who first approached me about writing a physics textbook. If it wasn't for his persuasion and his belief in me, I don't think I would have ever undertaken the writing of a textbook. Tim's suggestion to develop the art electronically also had a major impact on my approach to the development of the visual part of this book.

Albert Altman pointed out Ernst Mach's approach to developing mechanics starting with the law of conservation of momentum. Al encouraged me throughout the years as I struggled to reorganize the material around the conservation principles.

I am thankful to those who assisted on the first edition: Irene Nunes, who served as Development Editor through several iterations of the manuscript. Irene forced me to continuously rethink what I had written and her insights in physics kept surprising me. Her incessant questioning taught me that one doesn't need to be a science major to obtain a deep understanding of how the world around us works and that it is possible to explain physics in a way that makes sense for non-physics majors. Catherine Crouch helped write the final chapters of electricity and magnetism and the chapters on circuits and optics, permitting me to focus on the overall approach and the art program. Peter Dourmashkin helped me write the chapters on special relativity and thermodynamics. I appreciate Daryl Pedigo's hard work authoring, editing content, and coordinating the contributions to all of the practice material in this book. Also, Peter Dourmashkin and Wolfgang Rueckner helped me with some of the experiments that form the basis of the illustrations in a number of the chapters. Lastly, I would also like to thank my late uncle, Erich Lessing, for letting me use some of his beautiful pictures as chapter openers.

Many people provided feedback during the development of the first edition manuscript. I am particularly indebted to the late Ronald Newburgh and to Edward Ginsberg, who meticulously checked many of the chapters. I am also grateful to Edwin Taylor for his critical feedback on the special relativity chapter and to my colleague Gary Feldman for his suggestions for improving that chapter. Lisa Morris provided material for many of the Self-quizzes and my graduate students James Carey, Mark Winkler, and Ben Franta helped with data analysis and the appendices. The following people provided the material for the first edition Practice book: Wayne Anderson, Bill Ashmanskas, Linda Barton, Ronald Bieniek, Michael Boss, Anthony Buffa, Catherine Crouch, Peter Dourmashkin, Paul Draper, Andrew Duffy, Edward Ginsberg, William Hogan, Gerd Kortemeyer, Rafael Lopez-Mobilia, Chris-topher Porter, David Rosengrant, Gay Stewart, Christopher Watts, Lawrence Weinstein, Fred Wietfeldt, and Michael Wofsey.

I would also like to thank the editorial and production staff at Pearson. Margot Otway helped realize my vision for the art program. I would like to thank Darien Estes for helping conceive the second edition of this book, Jeanne Zalesky for her input, Deb Harden and Heidi Allgair for managing the process of getting things done, Karen Karlin who was my development editor for this edition, Jenny Moryan for coordinating the shooting of the videos, and Paul Corey for his continued support and belief in me. I am also indebted to David Bannon (Oregon State University), Jason Harlow (University of Toronto), and Michael Faux for their thorough accuracy checking. Frank Chmly helped prepare the reading questions for Mastering and Ken DeNisco helped with slides and Ready To Go Teaching Modules and with the accuracy checking of digital items. Finally, I would like to thank my colleague Greg Kestin for contributing his videos to the second edition.

I am also grateful to the participants of the NSF Faculty Development Conference "Teaching Physics Conservation Laws First" held in Cambridge, MA, in 1997. This conference helped validate and cement the approach in this book.

I am continuously indebted to the hundreds of students in Physics 1, Physics 11, and Applied Physics 50 who used early versions of this text in their course and provided the feedback that ended up turning my manuscript into a text that works not just for instructors but, more importantly, for students.

Reviewers of *Principles & Practice of Physics*

Over the years many people reviewed and class-tested the manuscript. The author and publisher are grateful for all of the feedback the reviewers provided, and we apologize if there are any names on this list that have been inadvertently omitted.

Edward Adelson, *Ohio State University*
Albert Altman, *University of Massachusetts, Lowell*
Susan Amador Kane, *Haverford College*
James Andrews, *Youngstown State University*
Arnold Arons, *University of Washington*
Robert Beichner, *North Carolina State University*
Bruce Birkett, *University of California, Berkeley*
David Branning, *Trinity College*
Bernard Chasan, *Boston University*
Stéphane Coutu, *Pennsylvania State University*
Corbin Covault, *Case Western Reserve University*
Catherine Crouch, *Swarthmore College*
Paul D'Alessandris, *Monroe Community College*
Paul Debevec, *University of Illinois at Urbana-Champaign*
N. John DiNardo, *Drexel University*
Margaret Dobrowolska-Furdyna, *Notre Dame University*
Paul Draper, *University of Texas, Arlington*
David Elmore, *Purdue University*
Robert Endorf, *University of Cincinnati*
Thomas Furtak, *Colorado School of Mines*
Ian Gatland, *Georgia Institute of Technology*
J. David Gavenda, *University of Texas, Austin*
Edward Ginsberg, *University of Massachusetts, Boston*
Gary Gladding, *University of Illinois*
Christopher Gould, *University of Southern California*
Victoria Greene, *Vanderbilt University*
Benjamin Grinstein, *University of California, San Diego*
Kenneth Hardy, *Florida International University*
Gregory Hassold, *Kettering University*
Peter Heller, *Brandeis University*
Laurent Hodges, *Iowa State University*
Mark Holtz, *Texas Tech University*
Zafar Ismail, *Daemen College*
Ramanathan Jambunathan, *University of Wisconsin Oshkosh*
Brad Johnson, *Western Washington University*
Dorina Kosztin, *University of Missouri Columbia*
Arthur Kovacs, *Rochester Institute of Technology* (deceased)
Dale Long, *Virginia Polytechnic Institute* (deceased)
John Lyon, *Dartmouth College*
Trecia Markes, *University of Nebraska, Kearney*
Peter Markowitz, *Florida International University*
Bruce Mason, *University of Oklahoma*
John McCullen, *University of Arizona*
James McGuire, *Tulane University*
Timothy McKay, *University of Michigan*
Carl Michal, *University of British Columbia*
Kimball Milton, *University of Oklahoma*
Charles Misner, *University of Maryland, College Park*
Sudipa Mitra-Kirtley, *Rose-Hulman Institute of Technology*
Delo Mook, *Dartmouth College*
Lisa Morris, *Washington State University*
Edmund Myers, *Florida State University*
Alan Nathan, *University of Illinois*
K.W. Nicholson, *Central Alabama Community College*
Fredrick Olness, *Southern Methodist University*
Dugan O'Neil, *Simon Fraser University*
Patrick Papin, *San Diego State University*
George Parker, *North Carolina State University*
Claude Penchina, *University of Massachusetts, Amherst*
William Pollard, *Valdosta State University*
Amy Pope, *Clemson University*
Joseph Priest, *Miami University* (deceased)
Joel Primack, *University of California, Santa Cruz*
Rex Ramsier, *University of Akron*
Steven Rauseo, *University of Pittsburgh*
Lawrence Rees, *Brigham Young University*
Carl Rotter, *West Virginia University*
Leonard Scarfone, *University of Vermont*
Michael Schatz, *Georgia Institute of Technology*
Cindy Schwarz, *Vassar College*
Hugh Scott, *Illinois Institute of Technology*
Janet Segar, *Creighton University*
Shahid Shaheen, *Florida State University*
David Sokoloff, *University of Oregon*
Gay Stewart, *University of Arkansas*
Roger Stockbauer, *Louisiana State University*
William Sturrus, *Youngstown State University*
Carl Tomizuka, *University of Illinois*
Mani Tripathi, *University of California–Davis*
Rebecca Trousil, *Skidmore College*
Christopher Watts, *Auburn University*
Robert Weidman, *Michigan Technological University*
Ranjith Wijesinghe, *Ball State University*
Augden Windelborn, *Northern Illinois University*

Acknowledgments for the Global Edition

Contributors

Mark Gino Aliperio, *De La Salle University*
Katarzyna Zuleta Estrugo, *École Polytechnique Fédérale de Lausanne*
Mukesh Kumar, *University of the Witwatersrand, Johannesburg*
Paul McKenna, *Glasgow Caledonian University*
Pankaj Mohanty, *University of Johannesburg*
Anouar Moustaj, *Utrecht University*
Abhishek Pandey, *University of the Witwatersrand, Johannesburg*
Wayne Rowlands, *Swinburne University of Technology*
Izak Snyman, *University of the Witwatersrand, Johannesburg*

Reviewers

Manoj Kumar Chauhan, *Jaypee Institute of Information Technology*
Joydeep Chowdhury, *Jadavpur University*
Ertuğrul Demir, *Yeditepe University*
Gülşen Görk, *Marmara University*
Bincy Susan Jacobs, *University of Johannesburg*
Paul McKenna, *Glasgow Caledonian University*
Iyabo Usman, *University of the Witwatersrand, Johannesburg*

Detailed Contents

About the Author 4
To the Student 5
To the Instructor 6
Acknowledgments 10

CHAPTER 1 **Foundations 19**

1.1 The scientific method 20
1.2 Symmetry 22
1.3 Matter and the universe 24
1.4 Time and change 26
1.5 Representations 27
1.6 Physical quantities and units 32
1.7 Significant digits 35
1.8 Solving problems 38
1.9 Developing a feel 42

CHAPTER 2 **Motion in One Dimension 53**

2.1 From reality to model 54
2.2 Position and displacement 55
2.3 Representing motion 57
2.4 Average speed and average velocity 59
2.5 Scalars and vectors 64
2.6 Position and displacement vectors 66
2.7 Velocity as a vector 70
2.8 Motion at constant velocity 71
2.9 Instantaneous velocity 73



CHAPTER 3 **Acceleration 86**

3.1 Changes in velocity 87
3.2 Acceleration due to gravity 88
3.3 Projectile motion 90
3.4 Motion diagrams 92
3.5 Motion with constant acceleration 96
3.6 Free-fall equations 99
3.7 Inclined planes 102
3.8 Instantaneous acceleration 103

CHAPTER 4 **Momentum 116**

4.1 Friction 117
4.2 Inertia 117
4.3 What determines inertia? 121
4.4 Systems 122
4.5 Inertial standard 127
4.6 Momentum 128
4.7 Isolated systems 130
4.8 Conservation of momentum 135

CHAPTER 5 **Energy 149**

5.1 Classification of collisions 150
5.2 Kinetic energy 151
5.3 Internal energy 153
5.4 Closed systems 156
5.5 Elastic collisions 160
5.6 Inelastic collisions 163
5.7 Conservation of energy 164
5.8 Explosive separations 166

CHAPTER 6 **Principle of Relativity 176**

6.1 Relativity of motion 177
6.2 Inertial reference frames 179
6.3 Principle of relativity 181
6.4 Zero-momentum reference frame 186
6.5 Galilean relativity 189
6.6 Center of mass 193
6.7 Convertible kinetic energy 198
6.8 Conservation laws and relativity 201

CHAPTER 7 **Interactions 211**

7.1 The effects of interactions 212
7.2 Potential energy 215
7.3 Energy dissipation 216
7.4 Source energy 219
7.5 Interaction range 222
7.6 Fundamental interactions 224
7.7 Interactions and accelerations 228
7.8 Nondissipative interactions 229
7.9 Potential energy near Earth's surface 232
7.10 Dissipative interactions 235

CHAPTER 8 **Force 246**

8.1 Momentum and force 247
8.2 The reciprocity of forces 248
8.3 Identifying forces 250
8.4 Translational equilibrium 251
8.5 Free-body diagrams 252
8.6 Springs and tension 255
8.7 Equation of motion 259
8.8 Force of gravity 262
8.9 Hooke's law 263
8.10 Impulse 265
8.11 Systems of two interacting objects 267
8.12 Systems of many interacting objects 269



CHAPTER 9 **Work 282**

9.1 Force displacement 283
9.2 Positive and negative work 284
9.3 Energy diagrams 286
9.4 Choice of system 289
9.5 Work done on a single particle 293
9.6 Work done on a many-particle system 296
9.7 Variable and distributed forces 300
9.8 Power 303

CHAPTER 10 **Motion in a Plane 314**

10.1 *Straight* is a relative term 315
10.2 Vectors in a plane 316
10.3 Decomposition of forces 319
10.4 Friction 322
10.5 Work and friction 323
10.6 Vector algebra 326
10.7 Projectile motion in two dimensions 328
10.8 Collisions and momentum in two dimensions 330
10.9 Work as the product of two vectors 331
10.10 Coefficients of friction 336

CHAPTER 11 **Motion in a Circle 351**

11.1 Circular motion at constant speed 352
11.2 Forces and circular motion 356
11.3 Rotational inertia 358
11.4 Rotational kinematics 362
11.5 Angular momentum 367
11.6 Rotational inertia of extended objects 372

CHAPTER 12 **Torque 388**

12.1 Torque and angular momentum 389
12.2 Free rotation 392
12.3 Extended free-body diagrams 393
12.4 The vectorial nature of rotation 395
12.5 Conservation of angular momentum 400
12.6 Rolling motion 404
12.7 Torque and energy 409
12.8 The vector product 411

CHAPTER 13 **Gravity 425**

13.1 Universal gravity 426
13.2 Gravity and angular momentum 431
13.3 Weight 434
13.4 Principle of equivalence 438
13.5 Gravitational constant 442
13.6 Gravitational potential energy 443
13.7 Celestial mechanics 446
13.8 Gravitational force exerted by a sphere 451



CHAPTER 14 **Special Relativity 461**

14.1 Time measurements 462
14.2 *Simultaneous* is a relative term 465
14.3 Space-time 469
14.4 Matter and energy 474

14.5 Time dilation 479
14.6 Length contraction 484
14.7 Conservation of momentum 488
14.8 Conservation of energy 492

CHAPTER 15 **Periodic Motion 504**

15.1 Periodic motion and energy 505
15.2 Simple harmonic motion 507
15.3 Fourier's theorem 509
15.4 Restoring forces in simple harmonic motion 511
15.5 Energy of a simple harmonic oscillator 516
15.6 Simple harmonic motion and springs 520
15.7 Restoring torques 524
15.8 Damped oscillations 527



CHAPTER 16 **Waves in One Dimension 539**

16.1 Representing waves graphically 540
16.2 Wave propagation 544
16.3 Superposition of waves 548
16.4 Boundary effects 550
16.5 Wave functions 556
16.6 Standing waves 562
16.7 Wave speed 564
16.8 Energy transport in waves 567
16.9 The wave equation 569

CHAPTER 17 **Waves in Two and Three Dimensions 580**

17.1 Wavefronts 581
17.2 Sound 584
17.3 Interference 587
17.4 Diffraction 592
17.5 Intensity 596
17.6 Beats 599
17.7 Doppler effect 602
17.8 Shock waves 607



CHAPTER 18 **Fluids 616**

18.1 Forces in a fluid 617
18.2 Buoyancy 622
18.3 Fluid flow 624
18.4 Surface effects 628
18.5 Pressure and gravity 635
18.6 Working with pressure 640
18.7 Bernoulli's equation 644
18.8 Viscosity and surface tension 647

CHAPTER 19 **Entropy 664**

19.1 States 665
19.2 Equipartition of energy 669
19.3 Equipartition of space 670
19.4 Evolution toward the most probable macrostate 673
19.5 Dependence of entropy on volume 678
19.6 Dependence of entropy on energy 683
19.7 Properties of a monatomic ideal gas 687
19.8 Entropy of a monatomic ideal gas 690

CHAPTER 20 **Energy Transferred Thermally 701**

20.1 Thermal interactions 702
20.2 Temperature measurement 706
20.3 Heat capacity 709
20.4 *PV* diagrams and processes 713
20.5 Change in energy and work 719
20.6 Isochoric and isentropic ideal gas processes 721
20.7 Isobaric and isothermal ideal gas processes 723

20.8 Entropy change in ideal gas processes 727
20.9 Entropy change in nonideal gas processes 731



CHAPTER 21 **Degradation of Energy 742**
21.1 Converting energy 743
21.2 Quality of energy 746
21.3 Heat engines and heat pumps 750
21.4 Thermodynamic cycles 756
21.5 Entropy constraints on energy transfers 761
21.6 Heat engine performance 764
21.7 Carnot cycle 768
21.8 Brayton cycle 770

CHAPTER 22 **Electric Interactions 783**
22.1 Static electricity 784
22.2 Electrical charge 785
22.3 Mobility of charge carriers 788
22.4 Charge polarization 793
22.5 Coulomb's law 796
22.6 Forces exerted by distributions of charge carriers 800

CHAPTER 23 **The Electric Field 813**
23.1 The field model 814
23.2 Electric field diagrams 816
23.3 Superposition of electric fields 817
23.4 Electric fields and forces 820
23.5 Electric field of a charged particle 824
23.6 Dipole field 825
23.7 Electric fields of continuous charge distributions 827
23.8 Dipoles in electric fields 832

CHAPTER 24 **Gauss's Law 845**
24.1 Electric field lines 846
24.2 Field line density 848
24.3 Closed surfaces 849
24.4 Symmetry and Gaussian surfaces 851
24.5 Charged conducting objects 854
24.6 Electric flux 858
24.7 Deriving Gauss's law 860
24.8 Applying Gauss's law 862



CHAPTER 25 **Work and Energy in Electrostatics 878**
25.1 Electric potential energy 879
25.2 Electrostatic work 880
25.3 Equipotentials 882
25.4 Calculating work and energy in electrostatics 886
25.5 Potential difference 889
25.6 Electrostatic potentials of continuous charge distributions 895
25.7 Obtaining the electric field from the potential 897



CHAPTER 26 **Charge Separation and Storage 908**

26.1 Charge separation 909
26.2 Capacitors 911
26.3 Dielectrics 915
26.4 Voltaic cells and batteries 917
26.5 Capacitance 921
26.6 Electric field energy and emf 925
26.7 Dielectric constant 928
26.8 Gauss's law in dielectrics 931

CHAPTER 27 **Magnetic Interactions 941**

27.1 Magnetism 942
27.2 Magnetic fields 944
27.3 Charge flow and magnetism 946
27.4 Magnetism and relativity 949
27.5 Current and magnetism 954
27.6 Magnetic flux 956
27.7 Moving particles in electric and magnetic fields 958
27.8 Magnetism and electricity unified 963

CHAPTER 28 **Magnetic Fields of Charged Particles in Motion 976**

28.1 Source of the magnetic field 977
28.2 Current loops and spin magnetism 978
28.3 Magnetic dipole moment and torque 980
28.4 Ampèrian paths 983
28.5 Ampère's law 987
28.6 Solenoids and toroids 990
28.7 Magnetic fields due to currents 993
28.8 Magnetic field of a moving charged particle 996

CHAPTER 29 **Changing Magnetic Fields 1011**

29.1 Moving conductors in magnetic fields 1012
29.2 Faraday's law 1014
29.3 Electric fields accompany changing magnetic fields 1015
29.4 Lenz's law 1016
29.5 Induced emf 1021
29.6 Electric field accompanying a changing magnetic field 1025
29.7 Inductance 1027
29.8 Magnetic energy 1029



CHAPTER 30 **Changing Electric Fields 1042**

30.1 Magnetic fields accompany changing electric fields 1043
30.2 Fields of moving charged particles 1046
30.3 Oscillating dipoles and antennas 1049
30.4 Displacement current 1055
30.5 Maxwell's equations 1059
30.6 Electromagnetic waves 1062
30.7 Electromagnetic energy 1067

CHAPTER 31 **Electric Circuits 1079**

31.1 The basic circuit 1080
31.2 Current and resistance 1082
31.3 Junctions and multiple loops 1084
31.4 Electric fields in conductors 1088
31.5 Resistance and Ohm's law 1092
31.6 Single-loop circuits 1096
31.7 Multiloop circuits 1101
31.8 Power in electric circuits 1106



CHAPTER 32 **Electronics 1121**

32.1 Alternating currents 1122
32.2 AC circuits 1124
32.3 Semiconductors 1128
32.4 Diodes, transistors, and logic gates 1130
32.5 Reactance 1136
32.6 *RC* and *RLC* series circuits 1141
32.7 Resonance 1148
32.8 Power in AC circuits 1150

CHAPTER 33 **Ray Optics 1164**

33.1 Rays 1165
33.2 Absorption, transmission, and reflection 1166
33.3 Refraction and dispersion 1169
33.4 Forming images 1173
33.5 Snel's law 1180
33.6 Thin lenses and optical instruments 1183
33.7 Spherical mirrors 1190
33.8 Lensmaker's formula 1193

CHAPTER 34 **Wave and Particle Optics 1207**

34.1 Diffraction of light 1208
34.2 Diffraction gratings 1210
34.3 X-ray diffraction 1213
34.4 Matter waves 1217
34.5 Photons 1218
34.6 Multiple-slit interference 1222
34.7 Thin-film interference 1227
34.8 Diffraction at a single-slit barrier 1229
34.9 Circular apertures and limits of resolution 1231
34.10 Photon energy and momentum 1234

Appendix A: Notation 1247
Appendix B: Mathematics Review 1257
Appendix C: SI Units, Useful Data, and Unit Conversion Factors 1263
Appendix D: Center of Mass of Extended Objects 1267
Appendix E: Derivation of the Lorentz Transformation Equations 1268
Solutions to Checkpoints 1271
Credits 1335
Index 1337



1 Foundations

1.1 The scientific method
1.2 Symmetry
1.3 Matter and the universe
1.4 Time and change
1.5 Representations

1.6 Physical quantities and units
1.7 Significant digits
1.8 Solving problems
1.9 Developing a feel

CONCEPTS

QUANTITATIVE TOOLS

Chances are you are taking this course in physics because someone told you to take it, and it may not be clear to you *why* you should be taking it. One good reason for taking a physics course is that, first and foremost, physics provides a fundamental understanding of the world. Furthermore, whether you are majoring in psychology, engineering, biology, physics, or something else, this course offers you an opportunity to sharpen your reasoning skills. Knowing physics means becoming a better problem solver (and I mean *real* problems, not textbook problems that have already been solved), and becoming a better problem solver is empowering: It allows you to step into unknown territory with more confidence. Before we embark on this exciting journey, let's map out the territory we are going to explore so that you know where we are going.

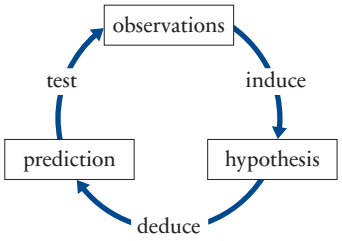
1.1 The scientific method

Physics, from the Greek word for "nature," is commonly defined as the study of matter and motion. Physics is about discovering the wonderfully simple unifying patterns that underlie absolutely everything that happens around us, from the scale of subatomic particles, to the microscopic world of DNA molecules and cells, to the cosmic scale of stars, galaxies, and planets. Physics deals with atoms and molecules; gases, solids, and liquids; everyday objects and black holes. Physics explores motion, light, and sound; the creation and annihilation of matter; evaporation and melting; electricity and magnetism. Physics is all around you: in the Sun that provides your daylight, in the structure of your bones, in your computer, in the motion of a ball you throw. In a sense, then, physics is the study of all there is in the universe. Indeed, biology, engineering, chemistry, astronomy, geology, and so many other disciplines you might name all use the principles of physics.

The many remarkable scientific accomplishments of ancient civilizations that survive to this day testify to the fact that curiosity about the world is part of human nature. Physics evolved from *natural philosophy*—a body of knowledge accumulated in ancient times in an attempt to explain the behavior of the universe through philosophical speculation—and became a distinct discipline during the scientific revolution that began in the 16th century. One of the main changes that occurred in that century was the development of the **scientific method,** an iterative process for going from observations to validated theories.

In its simplest form, the scientific method works as follows (**Figure 1.1**): A researcher makes a number of observations concerning either something happening in the natural world (a volcano erupting, for instance) or something happening during a laboratory experiment (a dropped brick and a dropped Styrofoam peanut travel to the floor at different speeds). These observations then lead the researcher to formulate a **hypothesis,** which is a tentative explanation of the observed phenomenon. The hypothesis is used to predict the outcome of some related natural occurrence (how a similarly shaped mountain near the erupting volcano will behave) or related laboratory experiment (what happens when a book and a sheet of paper are dropped at the same time). If the predictions prove inaccurate, the hypothesis must be modified. If the predictions prove accurate in test after test, the hypothesis is elevated to the status of either a **law** or a **theory.**

Figure 1.1 The scientific method is an iterative process in which a hypothesis, which is inferred from observations, is used to make a prediction, which is then tested by making new observations.



A law tells us *what* happens under certain circumstances. Laws are usually expressed in the form of relationships between observable quantities. A theory tells us *why* something happens and explains phenomena in terms of more basic processes and relationships. A scientific theory is not a mere conjecture or speculation. It is a thoroughly tested explanation of a natural phenomenon, one that is capable of making predictions that can be verified by experiment. The constant testing and retesting are what make the scientific method such a powerful tool for investigating the universe: The results obtained must be repeatable and verifiable by others.

CONCEPTS

EXERCISE 1.1 Hypothesis or not?

Which of the following statements are hypotheses? (*a*) Heavier objects fall to Earth faster than lighter ones. (*b*) The planet Mars is inhabited by invisible beings that are able to elude any type of observation. (*c*) Distant planets harbor forms of life. (*d*) Handling toads causes warts.

SOLUTION (*a*), (*c*), and (*d*). A hypothesis must be experimentally verifiable. (*a*) I can verify this statement by dropping a heavy object and a lighter one at the same instant and observing which one hits the ground first. (*b*) This statement asserts that the beings on Mars cannot be observed, which precludes any experimental verification and means this statement is not a valid hypothesis. (*c*) Although we humans currently have no means of exploring or closely observing distant planets, the statement is in principle testable. (*d*) Even though we know this statement is false, it *is* verifiable and therefore is a hypothesis.

Because of the constant reevaluation demanded by the scientific method, science is not a stale collection of facts but rather a living and changing body of knowledge. More

important, any theory or law *always* remains tentative, and the testing never ends. In other words, it is not possible to prove any scientific theory or law to be absolutely true (or even absolutely false). Thus the material you will learn in this book does not represent some "ultimate truth"—it is true only to the extent that it has not been proved wrong.

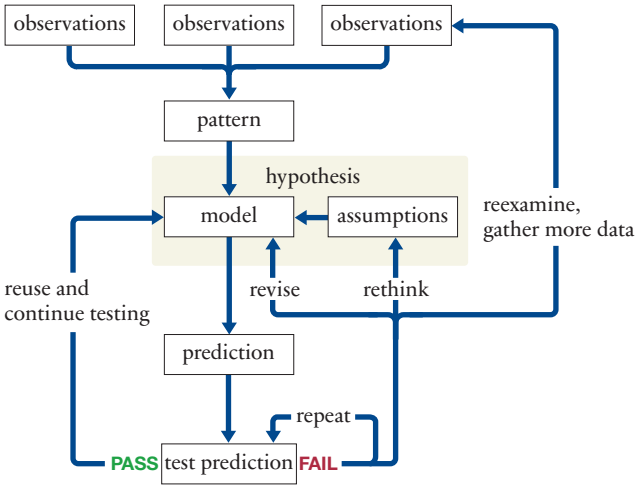
A case in point is *classical mechanics,* a theory developed in the 17th century to describe the motion of everyday objects (and the subject of most of this book). Although this theory produces accurate results for most everyday phenomena, from balls thrown in the air to satellites orbiting Earth, observations made during the last hundred years have revealed that under certain circumstances, significant deviations from this theory occur. It is now clear that classical mechanics is applicable for only a limited (albeit important) range of phenomena, and new branches of physics—*quantum mechanics* and the theory of *special relativity* among them—are needed to describe the phenomena that fall outside the range of classical mechanics.

The formulation of a hypothesis almost always involves developing a **model,** which is a simplified conceptual representation of some phenomenon. You don't have to be trained as a scientist to develop models. Everyone develops mental models of how people behave, how events unfold, and how things work. Without such models, we would not be able to understand our experiences, decide what actions to take, or handle unexpected experiences. Examples of models we use in everyday life are that door handles and door hinges are on opposite sides of doors and that the + button on a TV remote increases the volume or the channel number. In everyday life, we base our models on whatever knowledge we have, real or imagined, complete or incomplete. In science we must build models based on careful observation and determine ways to fill in any missing information.

Let's look at the iterative process of developing models and hypotheses in physics, with an eye toward determining what skills are needed and what pitfalls are to be avoided (**Figure 1.2**). Developing a scientific hypothesis often begins with recognizing patterns in a series of observations. Sometimes these observations are direct, but sometimes we must settle for indirect observations. (We cannot directly observe the nucleus of an atom, for instance, but a physicist can describe the structure of the nucleus and its behavior with great certainty and accuracy.) As Figure 1.2 indicates, the patterns that emerge from our observations must often be combined with simplifying assumptions to build a model. The combination of model and assumptions is what constitutes a hypothesis.

It may seem like a shaky proposition to build a hypothesis on assumptions that are accepted without proof, but making these assumptions—*consciously*—is a crucial step in making sense of the universe. All that is required is that, when formulating a hypothesis, we must be aware

Figure 1.2 Iterative process for developing a scientific hypothesis.



of these assumptions and be ready to revise or drop them if the predictions of our hypothesis are not validated. We should, in particular, watch out for what are called *hidden assumptions*—assumptions we make without being aware of them. As an example, try answering the following question. (Turn to the final section of this book, "Solutions to checkpoints," for the answer.)

1.1 I have two coins in my pocket, together worth 30 cents. If one of them is not a nickel, what coins are they?

Advertising agencies and magicians are masters at making us fall into the trap of hidden assumptions. Imagine a radio commercial for a new drug in which someone says, "Baroxan lowered my blood pressure tremendously." If you think that sounds good, you have made a number of assumptions without being aware of them—in other words, hidden assumptions. Who says, for instance, that lowering blood pressure "tremendously" is a good thing (dead people have tremendously low blood pressure) or that the speaker's blood pressure was too high to begin with?

Magic, too, involves hidden assumptions. The trick in some magic acts is to make you assume that something happens, often by planting a false assumption in your mind. A magician might ask, "How did I move the ball from here to there?" while in reality he is using two balls. I won't knowingly put false assumptions into *your* mind in this book, but on occasion you and I (or you and your instructor) may unknowingly make different assumptions during a given discussion, a situation that unavoidably leads to confusion and misunderstanding. Therefore it is important that we carefully analyze our thinking and watch for the assumptions that we build into our models.

If the prediction of a hypothesis fails to agree with observations made to test the hypothesis, there are several ways to address the discrepancy. One way is to rerun the

CONCEPTS

test to see if it is reproducible. If the test keeps producing the same result, it becomes necessary to revise the hypothesis, rethink the assumptions that went into it, or reexamine the original observations that led to the hypothesis.

EXERCISE 1.2 Stopped clock

A battery-operated wall clock no longer keeps time—neither hand moves. Develop a hypothesis explaining why it fails to work, and then make a prediction that permits you to test your hypothesis. Describe two possible outcomes of the test and what you conclude from the outcomes. (*Think* before you peek at the answer below.)

SOLUTION There are many reasons the clock might not run. Here is one example. Hypothesis: The batteries are dead. Prediction: If I replace the batteries with new ones, the clock should work. Possible outcomes: (1) The clock works once the new batteries are installed, which means the hypothesis is supported; (2) the clock doesn't work after the new batteries are installed, which means the hypothesis is not supported and must be either modified or discarded.

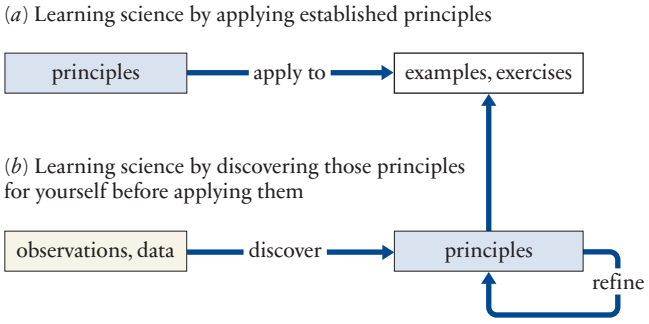
1.2 In Exercise 1.2, each of the conclusions drawn from the two possible outcomes contains a hidden assumption. What are the hidden assumptions?

The development of a scientific hypothesis is often more complicated than suggested by Figures 1.1 and 1.2. Hypotheses do not always start with observations; some are developed from incomplete information, vague ideas, assumptions, or even complete guesses. The refining process also has its limits. Each refinement adds complexity, and at some point the complexity outweighs the benefit of the increased accuracy. Because we like to think that the universe has an underlying simplicity, it might be better to scrap the hypothesis and start anew.

Figure 1.2 gives an idea of the skills that are useful in doing science: interpreting observations, recognizing patterns, making and recognizing assumptions, thinking logically, developing models, and using models to make predictions. It should not come as any surprise to you that many of these skills are useful in just about any context. Learning physics allows you to sharpen these skills in a very rigorous way. So, whether you become a financial analyst, a doctor, an engineer, or a research scientist (to name just a few possibilities), there is a good reason to take physics.

Figure 1.1 also shows that doing science—and physics in particular—involves two types of reasoning: *inductive*, which is arguing from the specific to the general, and *deductive*, arguing from the general to the specific. The most creative part of doing physics involves inductive reasoning, and this fact sheds light on how you might want to learn physics. One way, which is neither very useful nor very satisfying, is for me to simply tell you all the general principles physicists presently agree on and then for you to apply those principles in examples and exercises (**Figure 1.3*a***).This approach involves deductive reasoning only and robs you of the opportunity to learn the skill that is the most likely to benefit your career: discovering underlying patterns. Another way is for me to present you with data and observations and make you part of the discovery and refinement of the physics principles (Figure 1.3*b*). This approach is more time-consuming, and sometimes you may wonder why I'm not just *telling* you the final outcome. The reason is that discovery and refinement are at the heart of doing physics!

Figure 1.3



1.3 After reading this section, reflect on your goals for this course. Write down what you would like to accomplish and why you would like to accomplish this. Once you have done that, turn to the final section of this book, "Solutions to checkpoints," and compare what you have written with what I wrote.

1.2 Symmetry

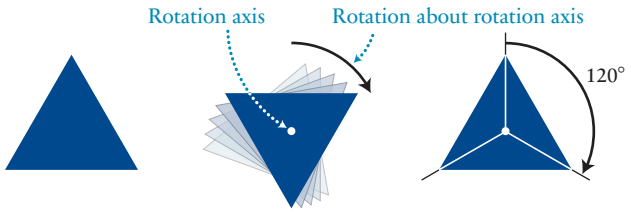
One of the basic requirements of any law of the universe involves what physicists call *symmetry,* a concept often associated with order, beauty, and harmony. We can define **symmetry** as follows: An object exhibits symmetry when certain operations can be performed on it without changing its appearance. Consider the equilateral triangle in **Figure 1.4*a***. If you close your eyes and someone rotates the triangle by 120° while you have your eyes closed, the triangle appears the same when you open your eyes, and you can't tell that it has been rotated. The triangle is said to have *rotational symmetry,* one of several types of geometrical symmetry.

Another common type of geometrical symmetry, *reflection symmetry,* occurs when one half of an object is the mirror image of the other half. The equilateral triangle in Figure 1.4 possesses reflection symmetry about the three axes shown in Figure 1.4*b*. If you imagine folding the triangle in half over each axis, you can see that the two halves are identical. Reflection symmetry occurs all around us: in the arrangement of atoms in crystals (**Figure 1.5*a*** and *b*) and in the anatomy of most life forms (Figure 1.5*c*), to name just two examples.

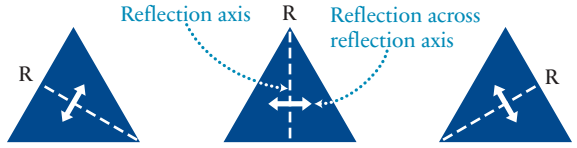
CONCEPTS

Figure 1.4

(*a*) Rotational symmetry: Rotating an equilateral triangle by 120° doesn't change how it looks



(*b*) Reflection symmetry: Across each reflection axis (labeled R), two sides of the triangle are mirror images of each other



The ideas of symmetry—that something appears unchanged under certain operations—apply not only to the shape of objects but also to the more abstract realm of physics. If there are things we can do to an experiment that leave the result of the experiment unchanged, then the phenomenon tested by the experiment is said to possess certain symmetries. Suppose we build an apparatus, carry out a certain measurement in a certain location, then move the apparatus to another location, repeat the measurement, and get the same result in both locations.* By moving the apparatus to a new location (*translating* it) and obtaining the same result, we have shown that the observed phenomenon has *translational symmetry*. Any physical law that describes this phenomenon must therefore mathematically exhibit translational symmetry; that is, the mathematical expression of this law must be independent of the location.

*In moving our apparatus, we must take care to move any relevant external influences along with it. For example, if Earth's gravity is of importance, then moving the apparatus to a location in space far from Earth does not yield the same result.

Likewise, we expect any measurements we make with our apparatus to be the same at a later time as at an earlier time; that is, translation in time has no effect on the measurements. The laws describing the phenomenon we are studying must therefore mathematically exhibit symmetry under translation in time; in other words, the mathematical expression of these laws must be independent of time.

EXERCISE 1.3 Change is no change

Figure 1.6 shows a snowflake. Does the snowflake have rotational symmetry? If yes, describe the ways in which the flake can be rotated without changing its appearance. Does it have reflection symmetry? If yes, describe the ways in which the flake can be split in two so that one half is the mirror image of the other.

Figure 1.6 Exercise 1.3.

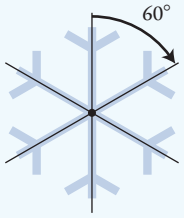


SOLUTION I can rotate the snowflake by 60° or a multiple of 60° (120°, 180°, 240°, 300°, and 360°) in the plane of the photograph without changing its appearance (**Figure 1.7*a***). It therefore has rotational symmetry.

I can also fold the flake in half along any of the three blue axes and along any of the three red axes in Figure 1.7*b*. The flake therefore has reflection symmetry about all six of these axes.

Figure 1.7

(*a*) Rotational symmetry



(*b*) Reflection symmetry

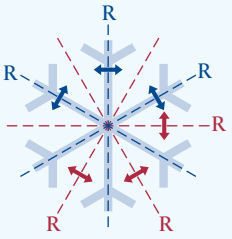
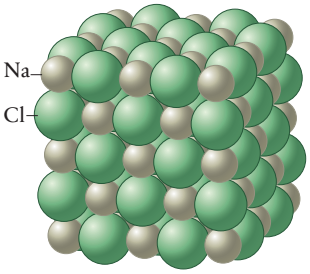


Figure 1.5 The symmetrical arrangement of atoms in a salt crystal gives these crystals their cubic shape.

(*a*) Micrograph of salt crystals



(*b*) Symmetrical arrangement of atoms in a salt crystal



(*c*) Da Vinci's Vitruvian Man shows the reflection symmetry of the human body



CONCEPTS

A number of such symmetries have been identified, and the basic laws that govern the inner workings of the physical world must reflect these symmetries. Some of these symmetries are familiar to you, such as translational symmetry in space or time. Others, like electrical charge or parity symmetry, are unfamiliar and surprising and go beyond the scope of this course. Whereas symmetry has always implicitly been applied to the description of the universe, it plays an increasingly important role in physics: In a sense the quest of physics in the 21st century is the search for (and test of) symmetries because these symmetries are the most fundamental principles that all physical laws must obey.

1.4 You always store your pencils in a cylindrical case. One day while traveling in the tropics, you discover that the cap, which you have placed back on the case day in, day out for years, doesn't fit over the case. What do you conclude?

1.3 Matter and the universe

The goal of physics is to describe all that happens in the universe. Simply put, the **universe** is the totality of matter and energy combined with the space and time in which all events happen—everything that is directly or indirectly observable. To describe the universe, we use *concepts*, which are ideas or general notions used to analyze natural phenomena.* To provide a quantitative description, these concepts must be expressed quantitatively, which requires defining a procedure for measuring them. Examples are the length or mass of an object, temperature, and time intervals. Such **physical quantities** are the cornerstones of physics. It is the accurate measurement of physical quantities that has led to the great discoveries of physics. Although many of the fundamental concepts we use in this book are familiar ones, quite a few are difficult to define in words, and we must often resort to defining these concepts in terms of the procedures used to measure them.

The fundamental physical quantity by which we map out the universe is **length**—a distance or an extent in space. The length of a straight or curved line is measured by comparing the length of the line with some standard length. In 1791, the French Academy of Sciences defined the standard unit for length, called the **meter** and abbreviated m, as one ten-millionth of the distance from the equator to the North Pole. For practical reasons, the standard was redefined in 1889 as the distance between two fine lines engraved on a bar of platinum-iridium alloy kept at the International Bureau of Weights and Measures near Paris. With the advent of lasers, however, it became possible to measure the speed of light with extraordinary accuracy, and so the meter was redefined in 1983 as the distance traveled by light in vacuum in a time interval of 1/299,792,458 of a second. This number is chosen so as to make the speed of light exactly 299,792,458 meters per second and yield a standard length for the meter that is very close to the length of the original platinum-iridium standard. This laser-based standard may never need to be revised.

1.5 Based on the early definition of the meter, one ten-millionth of the distance from the equator to the North Pole, what is Earth's radius?

Now that we have defined a standard for length, let us use this standard to discuss the structure and size scales of the universe. Because of the extraordinary range of size scales in the universe, we shall round off any values to the nearest power of ten. Such a value is called an **order of magnitude.** For example, any number between 0.3 and 3 has an order of magnitude of 1 because it is within a factor of 3 of 1; any number greater than 3 and equal to or less than 30 has an order of magnitude of 10. You determine the order of magnitude of any quantity by writing it in scientific notation and rounding the coefficient in front of the power of ten to 1 if it is equal to or less than 3 or to 10 if it is greater than 3.† For example, 3 minutes is 180 s, which can be written as 1.8×10^2 s. The coefficient, 1.8, rounds to 1, and so the order of magnitude is 1×10^2 s $= 10^2$ s. The quantity 680, to take another example, can be written as 6.8×10^2; the coefficient 6.8 rounds to 10, and so the order of magnitude is $10 \times 10^2 = 10^3$. And Earth's circumference is 40,000,000 m, which can be written as 4×10^7 m; the order of magnitude of this number is 10^8 m. You may think that using order-of-magnitude approximations is not very scientific because of the lack of accuracy, but the ability to work effectively with orders of magnitude is a key skill not just in science but also in any other quantitative field of endeavor.

All ordinary matter in the universe is made up of basic building blocks called **atoms** (Figure 1.8). Nearly all the matter in an atom is contained in a dense central nucleus, which consists of **protons** and **neutrons,** two types of subatomic particles. A tenuous *cloud* of **electrons,** a third type of subatomic particle, surrounds this nucleus. To an approximation, atoms are spherical and have a diameter of about 10^{-10} m. Spherical atomic nuclei have a diameter of about 10^{-15} m, making atoms mostly empty space. Atoms attract one another when they are a small distance apart but resist being squeezed into one another. The arrangement of atoms in a material helps determine the properties of the material.

CONCEPTS

*When an important concept is introduced in this book, the main word pertaining to the concept is printed in **boldface type.** All important concepts introduced in a chapter are listed at the back of this book, in the Glossary.

†The reason we use 3 in order-of-magnitude rounding, and not 5 as in ordinary rounding, is that orders of magnitude are logarithmic, and on this logarithmic scale log 3 = 0.48 lies nearly halfway between log 1 = 0 and log 10 = 1.

Figure 1.8 Scanning tunneling microscope image showing the individual atoms that make up a silicon surface. The size of each atom is about 1/50,000 the width of a human hair.

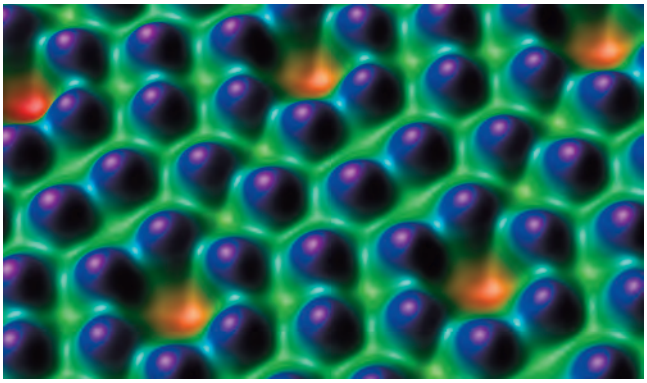


Figure 1.9 shows the relative size of some representative objects in the universe. The figure reveals a lot about the organization of matter in the universe and serves as a visual model of the structure of the universe. Roughly speaking, there is clustering of matter from smaller to larger at four length scales. At the subatomic scale, most of the matter in an atom is compressed into the tiny atomic nucleus, a cluster of subatomic particles. Atoms, in turn, cluster to form the objects and materials that surround us, from viruses to plants, animals, and other everyday objects. The next level is the clustering of matter in stars, some of which, such as the Sun, are surrounded by planets like Earth. Stars, in turn, cluster to form galaxies. As we shall discuss in Chapter 7, this clustering of matter reveals a great deal about the way different objects interact with one another.

EXERCISE 1.4 Tiny universe

If all the matter in the observable universe were squeezed together as tightly as the matter in the nucleus of an atom, what order of magnitude would the diameter of the universe be?

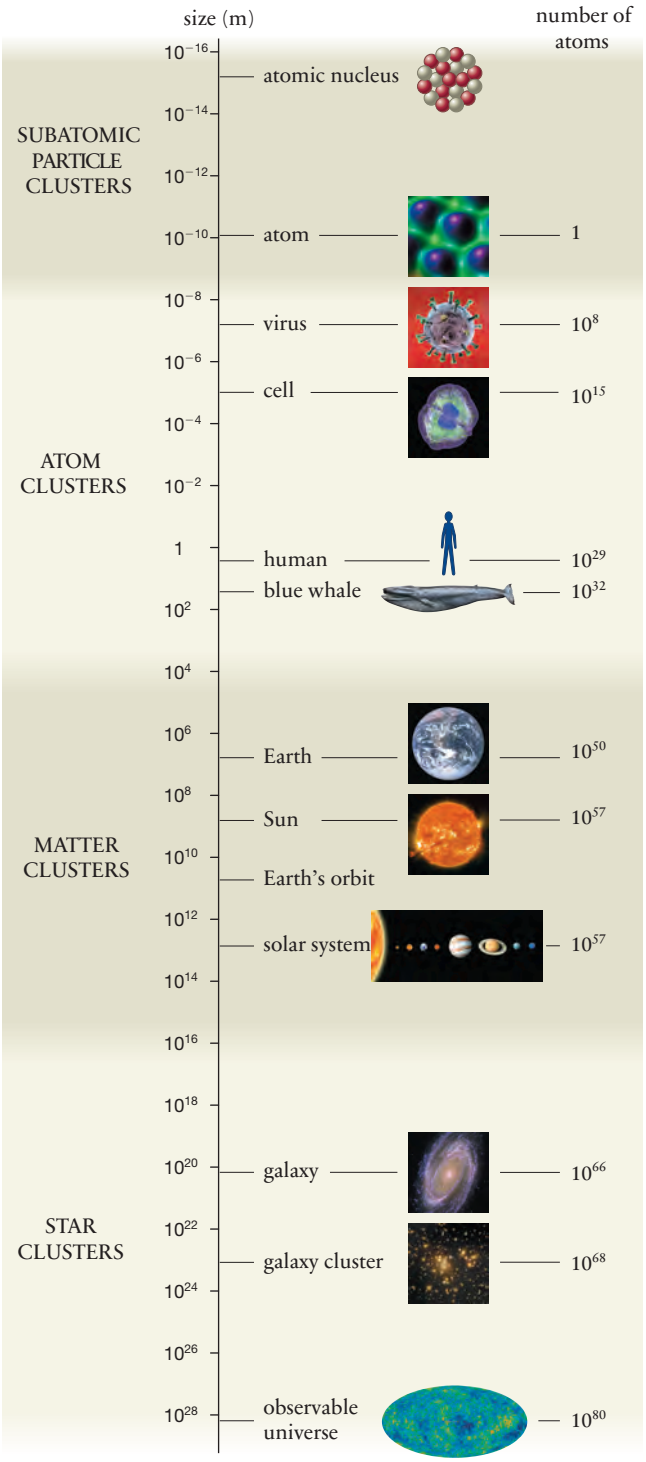
SOLUTION From Figure 1.9 I see that there are about 10^{80} atoms in the universe. I can arrange these atoms in a cube that has 10^{27} atoms on one side because such a cube could accommodate $10^{27} \times 10^{27} \times 10^{27} = 10^{81}$ atoms. Given that the diameter of a nucleus is about 10^{-15} m, the length of a side of this cube would be

$$(10^{27} \text{ atoms})(10^{-15} \text{ m per atom}) = 10^{12} \text{ m},$$

which is a bit larger than the diameter of Earth's orbit around the Sun.

An alternative method for obtaining the answer is to realize that the matter in a single nucleus occupies a cubic volume of about $(10^{-15} \text{ m})^3 = 10^{-45} \text{ m}^3$. If all the matter in the universe were squeezed together just as tightly, it would occupy a volume of about 10^{80} times the volume of an atomic nucleus, or $10^{80} \times 10^{-45} \text{ m}^3 = 10^{35} \text{ m}^3$. The side of a cube of this volume is equal to the cube root of 10^{35} m^3, or 4.6×10^{11} m, which is the same order of magnitude as my first answer.

Figure 1.9 A survey of the size and structure of the universe.



1.6 Imagine magnifying each atom in an apple to the size of the apple. What would the diameter of the apple then be?

CONCEPTS

1.4 Time and change

Profound and mysterious, time is perhaps the greatest enigma in physics. We all know what is meant by *time,* but it is difficult, if not impossible, to explain the idea in words. (Put the book down for a minute and try defining *time* in words before reading on.) One way to describe time is that it is the infinite continual progression of events in the past, present, and future, often experienced as a force that moves the world along. This definition is neither illuminating nor scientifically meaningful because it merely relates the concept of time to other, even less well-defined notions. Time is defined by the rhythm of life, by the passing of days, by the cycle of the seasons, by birth and death. However, even though many individual phenomena, such as the 24-hour cycle of the days, the cycle of seasons, and the swinging of a pendulum, are repetitive, the time we experience does not appear to be repetitive, and in classical physics time is considered to be a continuous succession of events.

The irreversible flow of time controls our lives, pushing us inexorably forward from the past to the future. Whereas we can freely choose our location and direction in all three dimensions in space, time flows in a single direction, dragging us forward with it. Time thus presents less symmetry than the three dimensions of space: Although opposite directions in space are equivalent, opposite directions in time are not equivalent. The "arrow of time" points only into the future, a direction we define as the one we have no memory of. Curiously, most of the laws of physics have no requirement that time has to flow in one direction only, and it is not until Chapter 19 that we can begin to understand why events in time are essentially irreversible.

The arrow of time allows us to establish a *causal relationship* between events. For example, lightning causes thunder and so lightning has to occur *before* the thunder. This statement is true for all observers: No matter who is watching the storm and no matter where that storm is happening, every observer first sees a lightning bolt and only after that hears the thunder because an effect never precedes its cause. Indeed, the very organization of our thoughts depends on the **principle of causality:**

> **Whenever an event A causes an event B, all observers see event A happening first.**

Without this principle, it wouldn't be possible to develop any scientific understanding of how the world works. (No physics course to take!) The principle of causality also makes it possible to state a definition: **Time** is a physical quantity that allows us to determine the sequence in which related events occur.

To apply the principle of causality and sort out causes and effects, it is necessary to develop devices—clocks—for keeping track of time. All clocks operate on the same principle: They repeatedly return to the same state. The rotation of Earth about its axis can serve as a clock if we note the instant the Sun reaches its highest position in the sky on successive days. A swinging pendulum, which repeatedly returns to the same vertical position, can also serve as a clock. The time interval between two events can be determined by counting the number of pendulum swings between the events. The accuracy of time measurements can be increased by using a clock that has frequent repetition of events.



1.7 (*a*) State a possible cause for the following events: (*i*) The light goes out in your room; (*ii*) you hear a loud, rumbling noise; (*iii*) your debit card is rejected at a store. (*b*) Could any of the causes you named have occurred after their associated event? (*c*) Describe how you feel when you experience an event but don't know what caused it—you hear a strange noise when camping, for instance, or an unexpected package is sitting on your doorstep.

The familiar standard unit for measuring time is the **second** (abbreviated s), originally defined as 1/86,400 of a day but currently more accurately defined as the duration of 9,192,631,770 periods of certain radiation emitted by cesium atoms. **Figure 1.10** gives an idea of the vast range of time scales in the universe.

The English physicist Isaac Newton stated, "Absolute, true, and mathematical time, of itself and from its own nature, flows equably without relation to anything external." In other words, the notion of past, present, and future is universal—"now" for you, wherever you are at this instant, is also "now" everywhere else in the universe. Although this notion of the universality of time, which is given the name **absolute time,** is intuitive, experiments described in Chapter 14 have shown this notion to be false. Still, for many experiments and for the material we discuss in most of this book, the notion of absolute time remains an excellent approximation.

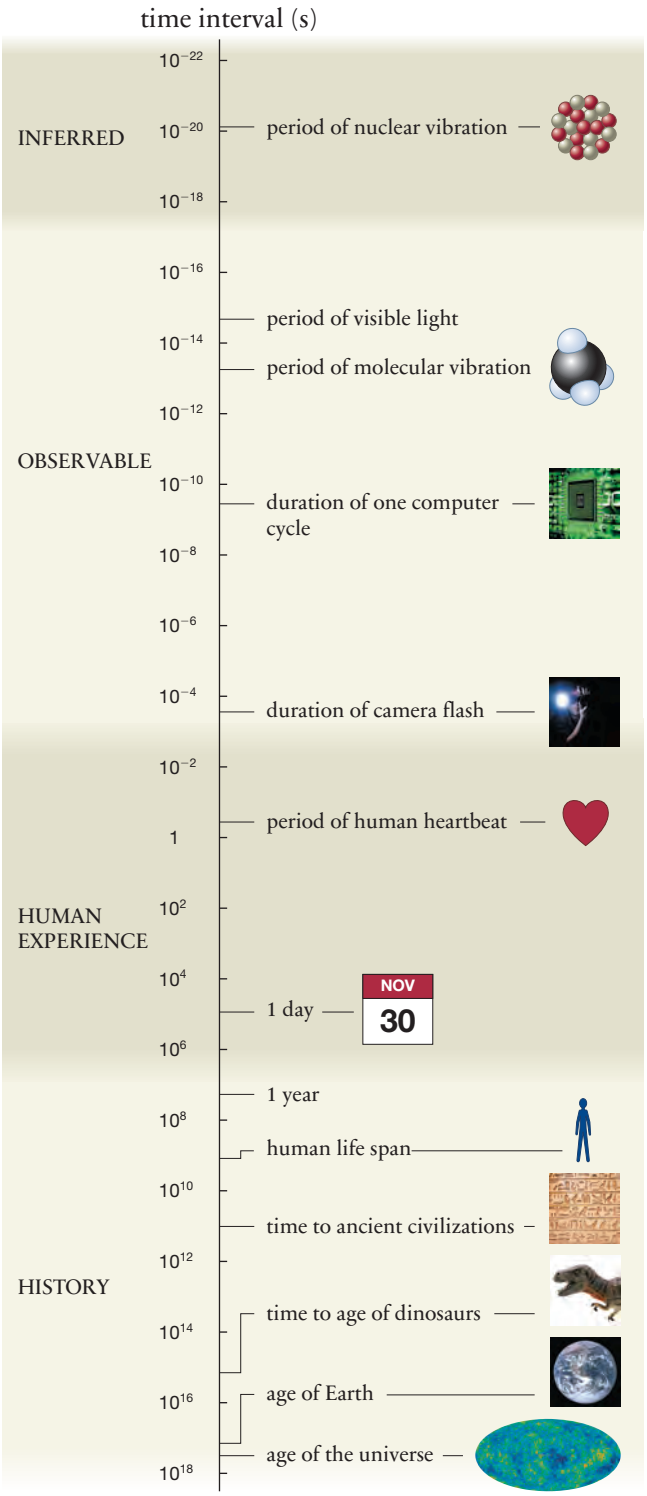
Now that we have introduced space and time, we can use these concepts to study events. Throughout this book, we focus on **change,** the transition from one state to another. Examples of change are the melting of an ice cube, motion (a change in location), the expansion of a piece of metal, the flow of a liquid. As you will see, one might well call physics the study of the changes that surround us and convey the passage of time. What is most remarkable about all this is that we shall discover that underneath all the changes we'll look at, certain properties remain *unchanged*. These properties give rise to what are called *conservation laws,* the most fundamental and universal laws of physics.

There is a profound aesthetic appeal in knowing that symmetry and conservation are the cornerstones of the laws that govern the universe. It is reassuring to know that an elegant simplicity underlies the structure of the universe and the relationship between space and time.

CONCEPTS

1.8 A single chemical reaction takes about 10^{-13} s. What order of magnitude is the number of sequential chemical reactions that could take place during a physics class?

Figure 1.10 A survey of the time scales on which events in the universe take place.



1.5 Representations

Of all our senses, vision may be the one that most informs our mind. For this reason, expert problem solvers rarely start working on a problem without first making some sort of visual representation of the available information, and you should always do the same. Such a drawing helps you establish a clear mental image of the situation, interpret the problem in light of your own knowledge and experience, develop a qualitative understanding of the problem, and organize the information in a meaningful way. Without the drawing, you have to juggle all the information in your head. The drawing frees your brain to deal with the solution. As an example, consider the situation described in Checkpoint 1.9. First try solving the problem without a piece of paper and note the mental effort it takes; then make a sketch and work out the solution.

1.9 You and your spouse are working out a seating arrangement around a circular table for dinner with John and Mary Jones, Mike and Sylvia Masters, and Bob and Cyndi Ahlers. Mike is not fond of the Ahlers, and Sylvia asked that she not be seated next to John. You would like to alternate men and women and avoid seating spouses next to each other. Determine an arrangement that satisfies all the constraints.

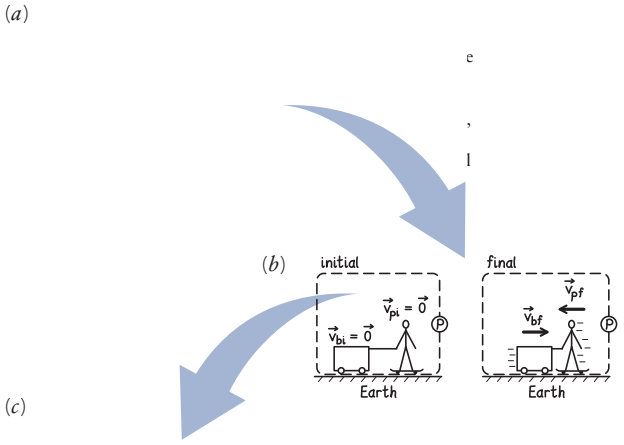
Were you able to solve the problem without a sketch? Doing so would be difficult because there is more information than most people can comfortably keep in mind at once. When you represent the information visually, however, it is relatively easy to solve this problem (see the solution in the back of this book). The drawing breaks the problem down into small steps and helps you articulate in your mind what you are trying to accomplish.

Visual representations are not helpful only for making seating arrangements. They work just as well for solving physics problems, although it may not immediately be clear *how* to represent the available information. For this reason we shall develop a number of different ways—pictures, sketches, diagrams, graphs—and a number of different context-specific procedures to represent information visually in our study of physics. As you will see, these visual representations are an integral part of getting a grip on a problem and developing a model (**Figure 1.11** on the next page).

As we discussed in Section 1.1, descriptions of the physical world always begin with simplified representations. When you solve physics problems, elaborate drawings clutter your mind with irrelevant information and prevent you from getting a clear view of the important features. One of the most basic skills in physics, therefore, is to decide what to leave out of your drawings. If you leave out essential features, the representation is useless, but if you put in too many details, it becomes impossible to analyze the situation.

CONCEPTS

Figure 1.11 Multiple representations help you solve problems. (*a*) Many problems start with a verbal representation. (*b*) Turning the words into a sketch helps you to grasp the problem. (*c*) The sketch can then give meaning to a mathematical representation of the problem.





Sometimes it is necessary to begin with an oversimplification in order to develop a feel for a given situation. Once this initial understanding has been gained, however, it becomes possible to construct less idealized models, with each successive model being a more realistic representation of the real-world situation.

In any drawings you make, therefore, you should treat everyday objects as simplifications that can be characterized by a minimum number of features or quantities. Some joke that physicists hold a grossly oversimplified view of the world, thinking in such terms as, for instance, "Consider a spherical cow. . . ." The world around us, cows included, contains infinitely many details that may play a role in the grand scheme of things, but to get a grip on any problem it is important to begin by leaving out as much detail as possible. If the resulting model reproduces the main features of the real world, you know you have taken the essential attributes into account. As an example, **Figure 1.12** shows a progression from a photograph of a cow to an abstract rendition of it. To study the pattern on the cow's hide, you need the photograph. If you are interested in only the position of a certain cow in a certain pasture, however, a simple dot suffices (Figure 1.12*d*). By reducing the cow to a dot, you have discarded any information about its size and shape, but as you will see as you continue with your study of physics, this information is often not relevant.

EXERCISE 1.5 Stretching a spring

For a physics laboratory assignment, one end of a spring is attached to a horizontal rod so that the spring hangs vertically, and a ruler is hung vertically alongside the spring. The stretching properties of the spring are to be measured by attaching eight identical beads to the spring's free end. With no beads attached, the free end of the spring is at a ruler reading of 23.4 mm. With one bead attached, the end of the spring drops to 25.2 mm. When the second, third, and fourth beads are attached one at a time, the end drops to ruler readings of 26.5 mm, 29.1 mm, and 30.8 mm, respectively. Adding the fifth and sixth beads together moves the spring end to 34.3 mm, and adding the last two beads moves the end to 38.2 mm. (*a*) Make a pictorial representation of this setup. (*b*) Tabulate the data. (*c*) Plot the data on a graph, showing the ruler readings on the vertical axis and the numbers of beads on the horizontal axis. (*d*) Describe what can be inferred from the data.

SOLUTION (*a*) The important items that should appear in my drawing are the spring, the rod from which it is suspended, the ruler, and at least one bead. My drawing should also indicate how the ruler readings are obtained. I can illustrate this procedure with just one bead, as **Figure 1.13*a*** shows. By drawing the spring first with no beads attached and then with one bead attached and showing the two ruler readings, I've represented the general procedure of adding beads one (or two) at a time and paying attention to how each addition changes the position of the spring end.

(*b*) See Figure 1.13*b*.

(*c*) See Figure 1.13*c*.

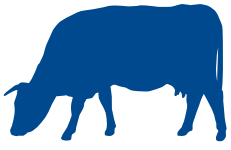
(*d*) The relationship between the ruler readings and the numbers of suspended beads is linear. That is to say, each additional bead stretches the spring by about the same amount.

Figure 1.12 Increasingly simplified and abstract representations of a cow.

(*a*) Photograph



(*b*) Simplified sketch



(*c*) Rectangle: represents cow's position and extent



(*d*) Dot: shows only cow's position

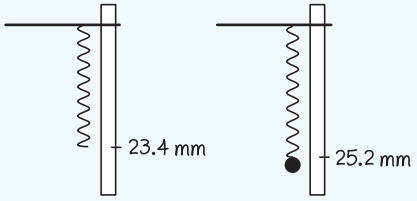




CONCEPTS

Figure 1.13

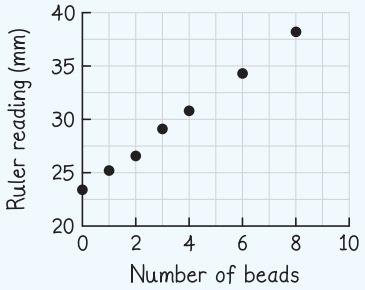
(a) Sketch



(b) Table

Number of beads	Ruler reading (mm)
0	23.4
1	25.2
2	26.5
3	29.1
4	30.8
6	34.3
8	38.2

(c) Graph



Exercise 1.5 demonstrates how various ways of representing information help you interpret data. The pictorial representation of Figure 1.13*a* helps you visualize how the measurements were taken. The table of Figure 1.13*b* organizes the data, and the graph of Figure 1.13*c* allows you to recognize the linear relationship between how far the spring stretches and the number of beads suspended from it—something that is not at all obvious from the text or the table.

Like the sketch, the graph is a simplified representation of the stretching of the spring. Each representation involves a loss of information and detail. The sketch is a simplified two-dimensional representation of a three-dimensional setup, and the graph shows only one piece of information for each measurement: the position of the bottom of the spring. All other information is left out in order to reveal one crucial point: How much the spring stretches is proportional to the number of beads suspended from it (something we look at in more detail in Chapter 9).

You may be beginning to wonder what the role of mathematics is in physics, given that we haven't used any thus far. One of the main roles of mathematics in physics is allowing us to express succinctly and unambiguously ideas that if expressed verbally would require a lot of words whose meaning may not be precise. Take this statement, for example:

> **The magnitude of the acceleration of an object is directly proportional to the magnitude of the vector sum of the forces exerted on the object and inversely proportional to the object's inertia. The direction of the acceleration is the same as the direction of the vector sum of the forces.**

(Don't worry about the meaning of this statement for now; just note that it is quite a mouthful.) We can express this statement mathematically as

$$\vec{a} = \frac{\Sigma \vec{F}}{m}.$$

These two expressions—one verbal, the other mathematical—are identical. Neither one is more accurate than the other. Right now the symbols in the equation may make no more sense to you than hieroglyphs. Once you understand their meaning, however, the mathematical expression can be parsed much more quickly than the verbal expression can. In this respect, mathematics plays the same role as visual representations: It relieves the brain of having to keep track of many words. Another important benefit of mathematical representation is that we can use the techniques of mathematics to manipulate the symbols and obtain new insights.

Without an understanding of the meaning of the concepts in any expression (acceleration and force in our example here), however, verbal expressions and mathematical ones are both meaningless, and so it is important to focus first on the meanings of concepts. As you will notice, this book has been designed to develop concepts first and to emphasize the visual representation of these concepts before moving on to a mathematical treatment (see the box "Organization of this book," next page).

CONCEPTS

1.10 Picture a long, straight corridor running east-west, with a water fountain located somewhere along it. Starting from the west end of the corridor, a woman walks a short distance east along the corridor and stops before reaching the water fountain. The distance from her to the fountain is twice the distance she has walked. She then continues walking east, passes the water fountain, and stops 60 m from her first stop. Now the distance from her to the fountain is twice the distance from her to the east end of the corridor. How long is the corridor?

Organization of this book

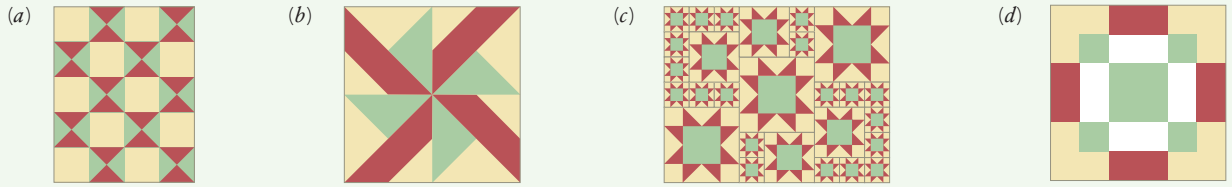
Each chapter is divided into two parts: *Concepts* and *Quantitative Tools*. The *Concepts* part develops the conceptual framework for the subject of the chapter, concentrating on the underlying ideas and helping you develop a mental picture of the subject. The *Quantitative Tools* part develops the mathematical framework, building on the ideas developed in the *Concepts* part. Interspersed in the text of each chapter are several types of learning aids to guide you through the chapter and a variety of questions and problems that allow you to apply and sharpen your understanding of physics:

1. **Checkpoints** (). These questions compel you to test yourself on how well you understand the material you just read. Do not skip them as you read the text. First of all, the answer to a checkpoint may be necessary to understand the text following that checkpoint. Second, and more important, I've put these checkpoints right in the text because working on them means learning the material. The answers to all the checkpoints are at the end of this book.
2. **Exercises and examples.** The fully-worked-out exercises and examples help you develop and apply problem-solving strategies. It is generally a good idea to attempt to solve the problem by yourself before reading the solution. More information on general approaches to problem solving is given in Section 1.8.
3. **Procedures.** Approaches for analyzing specific situations are given in separate, highlighted boxes.
4. **Self-quiz.** The Self-quiz in each chapter, which always comes at the end of the conceptual part of the chapter, allows you to assess your understanding of the concepts before you move on to the quantitative treatment. Complete each Self-quiz before working on the quantitative part of the chapter, even if you are already familiar with most of the material covered. Before tackling the quantitative material, be sure to resolve any difficulties you might have in answering a Self-quiz question, either by rereading the material in the conceptual part of the chapter or by consulting your instructor.
5. **Chapter summary.** The summary is just what the name implies, a condensed record of the key elements from the chapter.
6. **Questions and problems.** After working your way to this point, you should be ready to solve some problems on your own. These problems are both conceptual and quantitative, organized by chapter section and labeled with a rough "degree of difficulty" scale, indicated by one, two, or three dots. One-dot problems are fairly straightforward and usually involve only one major concept. Two-dot problems typically require you to put together two or more ideas from the chapter, or even to combine some element of the current chapter with material from other chapters. Three-dot problems are more challenging, or even a bit tricky. Some of these are designated as "CR" (context-rich), a category that is described later in this chapter. Context-rich problems are in the **Additional Problems** at the end of this section, together with other general problems.
7. **Glossary.** At the end of this book is a list defining the important concepts in each chapter (which are the terms printed in bold).

Self-quiz

1. Two children in a playground swing on two swings of unequal length. The child on the shorter swing is considerably heavier than the child on the longer swing. You observe that the longer swing swings more slowly. Formulate a hypothesis that could explain your observation. How could you test your hypothesis?
2. What symmetries do you observe in the quilt patterns of **Figure 1.14**?

Figure 1.14

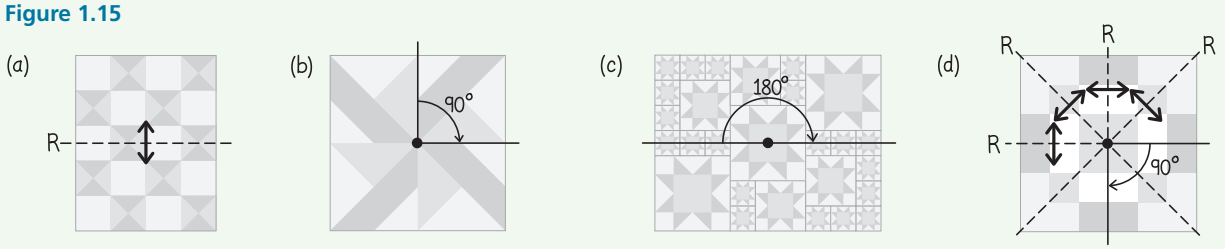


3. Give the order of magnitude of these quantities in meters or seconds: (*a*) length of a football field, (*b*) height of a mature tree, (*c*) one week, (*d*) one year.
4. Starting from the first floor, an elevator stops at floors 5, 2, 4, 3, 6, and 4 (in that order) before returning to the first floor. (*a*) Represent this motion visually. (*b*) If the floors are evenly spaced and the distance between floors 1 and 6 is 15 m, what is the total distance traveled by the elevator?

Answers

1. One hypothesis is that longer swings swing more slowly than shorter swings. You can test this hypothesis by adjusting the length of either swing until the two lengths are the same and then asking the children to remount their respective swings and swing again. If the originally longer swing is still the slower one, your hypothesis is not correct. If the two swings now have the same speed, your hypothesis is correct. Another hypothesis is that heavier children swing faster than lighter ones. You can test this hypothesis by asking the children to trade places. If the longer swing now swings faster than the shorter swing, your hypothesis is correct. If the longer swing still swings more slowly, your hypothesis is incorrect.
2. See **Figure 1.15**. (*a*) Reflection symmetry about a horizontal line through the center. (*b*) Rotational symmetry by multiples of 90°. (*c*) Rotational symmetry by 180°. (*d*) Rotational symmetry by multiples of 90° and reflection symmetry about a horizontal, vertical, or diagonal line through the center.

Figure 1.15



3. (*a*) 100 yards is about 100 m; the order of magnitude is thus 100 m $= 10^2$ m. (*b*) An average mature tree is between 5 and 20 m tall, for an average of 12 m $= 1.2 \times 10^1$ m. The coefficient 1.2 rounds to 1, and so the order of magnitude is 1×10^1 m $= 10$ m. (*c*) 1 week $=$ (1 week)(7 days/week)(24 h/day) $\cdot$ (60 min/h)(60 s/min) $= 604{,}800$ s $= 6 \times 10^5$ s; the coefficient 6 rounds to 10, and so the order of magnitude is 10×10^5 s $= 10^6$ s. (*d*) 1 year $= 52$ weeks $=$ (52 weeks)(604,800 s/week) $= 31{,}449{,}600$ s $= 3.1 \times 10^7$ s; the coefficient 3.1 rounds to 10, and so the order of magnitude is 10×10^7 s $= 10^8$ s.
4. (*a*) See **Figure 1.16** for one way to represent the motion. Note that the elevator itself is not represented because showing it would add nothing we need to the visual information. The only thing we are interested in is distances traveled, represented by the vertical lines. (*b*) If the distance between floors 1 and 6 is 15 m, one floor is (15 m)/5 $= 3.0$ m. From the figure, I see that the numbers of floors traveled between successive stops are 4, 3, 2, 1, 3, 2, and 3, for a total of 18 floors, or 18(3.0 m) $= 54$ m.

Figure 1.16

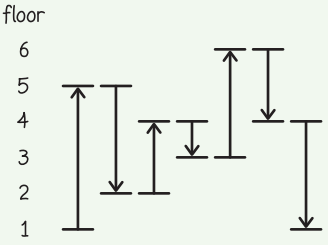


Table 1.1 Physical quantities and their symbols

Physical quantity	Symbol
length	ℓ
time	t
mass	m
speed	v
volume	V
energy	E
temperature	T

1.6 Physical quantities and units

Because physics is a quantitative science, statements must be expressed in numbers, which requires either measuring or calculating numerical values for physical quantities. In this section we review some basic rules for dealing with physical quantities, which in this book are represented by italic symbols—typically letters from the Roman or Greek alphabet, such as t for time and σ for electrical conductivity. Table 1.1 gives the symbols for some of the physical quantities we use throughout the book.

Physical quantities are expressed as the product of a numerical value and a unit of measurement. For example, the length ℓ of an object that is 1.2 m long can be expressed as $\ell = 1.2$ m. The unit system used in science and engineering throughout the world and in everyday life in most countries is the Système International (International System), and the units are collectively called **SI units.** This system consists of seven base units (Table 1.2) from which all other units currently in use can be derived. For example, the physical quantity speed, which we discuss in Chapter 2, is defined as the distance traveled divided by the time interval over which the travel takes place. Thus the SI derived unit of speed is meters per second (m/s), the base unit of length divided by the base unit of time. A list of SI derived units and their relationship to the seven base units is given in Appendix C.

Be careful not to confuse abbreviations for units with symbols for physical quantities. Unit abbreviations are printed in roman (upright) type—m for meters, for instance—and symbols for physical quantities are printed in italic (slanted) type—t for time, say. Also, bear in mind that you can add and subtract quantities only if they have the same units; it is meaningless to add, say, 3 m to 4 kg.

To produce multiples of any SI unit and conveniently work with very large or very small numbers, we modify the unit name with prefixes representing integer powers of ten (Table 1.3). For example, a billionth of a second is denoted by 1 ns (pronounced "one nanosecond"):

$$1 \text{ ns} = 10^{-9} \text{ s}. \tag{1.1}$$

One thousand meters is denoted by 1 km, "one kilometer." Prefixes are never used without a unit and are never combined into compound prefixes. The unit *kilogram* contains a prefix (*kilo-*) because it is derived from the non-SI unit *gram* (1 kg = 1000 g). Therefore 10^{-6} kg never becomes 1 μkg. Instead, the names and multiples of the kilogram are constructed by adding the appropriate prefix to the word *gram* and the abbreviation g. For example, 10^{-6} kg becomes 1 mg, "one milligram."

The standard practice in engineering is to use only the powers of ten that are multiples of three.

Table 1.2 The seven SI base units

Name of unit	Abbreviation	Physical quantity
meter	m	length
kilogram	kg	mass
second	s	time
ampere	A	electric current
kelvin	K	thermodynamic temperature
mole	mol	amount of substance
candela	cd	luminous intensity

Table 1.3 SI prefixes

10^n	Prefix	Abbreviation	10^n	Prefix	Abbreviation
10^0	—	—			
10^3	kilo-	k	10^{-3}	milli-	m
10^6	mega-	M	10^{-6}	micro-	μ
10^9	giga-	G	10^{-9}	nano-	n
10^{12}	tera-	T	10^{-12}	pico-	p
10^{15}	peta-	P	10^{-15}	femto-	f
10^{18}	exa-	E	10^{-18}	atto-	a
10^{21}	zetta-	Z	10^{-21}	zepto-	z
10^{24}	yotta-	Y	10^{-24}	yocto-	y



1.11 Use prefixes from Table 1.3 to remove either all or almost all of the zeros in each expression. (*a*) $\ell = 150{,}000{,}000$ m, (*b*) $t = 0.000\ 000\ 000\ 012$ s, (*c*) 1200 km/s, (*d*) 2300 kg.

Of the seven SI base units, we have already discussed two, the meter and the second. We discuss the base unit for mass, the kilogram, in Chapter 4, the base unit for electric current, the ampere, in Chapter 27, and the base unit for temperature, the kelvin, in Chapter 20.

The **mole** (abbreviated mol) is the SI base unit that measures the quantity of a given substance. A mole is approximately the number of atoms in 12×10^{-3} kg of carbon-12, the most common form of carbon. This number is called **Avogadro's number** N_A, after the 19th-century Italian physicist Amedeo Avogadro. In the SI system, Avogadro's number is defined to be exactly

$$N_A \equiv 6.02214076 \times 10^{23}. \tag{1.2}$$

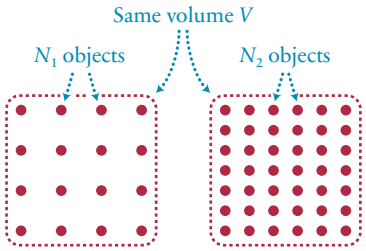
Note that the mole is simply a number: Just as *one dozen* means 12 of anything and *one gross* means 144 of anything, *one mole* means 6.022×10^{23} of anything. So 1 mol of helium atoms is 6.022×10^{23} helium atoms, and 1 mol of carbon dioxide molecules is 6.022×10^{23} carbon dioxide molecules.

The final SI base unit, the *candela,* measures luminous intensity. One candela (1 cd) is roughly the amount of light generated by a single candle; the light emitted by a 100-watt light bulb is about 120 cd. The definition of the candela takes into account how the human eye perceives the intensity of various colors and is therefore rather unwieldy. For this reason we do not use the candela in this book in the chapters dealing with light, concentrating instead on the amount of energy carried by light.

An important concept used throughout physics is **density,** the physical quantity that measures how much of some substance there is in a given volume. Depending on the quantity being measured, there are various types of density. For example, *number density* is the number of objects per unit volume. If there are N objects in a volume V, then the number density n of these objects is

$$n \equiv \frac{N}{V}. \tag{1.3}$$

Figure 1.17 Number density.



The greater the number N of objects in a given space V, the higher the number density $n = N/V$. In this case $N_2 > N_1$, so $n_2 > n_1$.

(The symbol $\equiv$ means that the equality is either a definition or a convention.) If the objects in a given volume are packed together more tightly, the number density is higher (**Figure 1.17**). *Mass density* ρ (Greek rho) is the amount of mass m per unit volume:

$$\rho \equiv \frac{m}{V}. \tag{1.4}$$

EXERCISE 1.6 Helium density

At room temperature and atmospheric pressure, 1 mol of helium gas has a volume of 24.5×10^{-3} m^3. The same amount of *liquid* helium has a volume of 32.0×10^{-6} m^3. What are the number and mass densities of (*a*) the gaseous helium and (*b*) the liquid helium? The mass of one helium atom is 6.647×10^{-27} kg.

SOLUTION (*a*) I know from Eq. 1.2 that 1 mol of helium contains 6.022×10^{23} atoms, and I can use this information in Eq. 1.3 to get the number density:

$$n = \frac{6.022 \times 10^{23} \text{ atoms}}{24.5 \times 10^{-3} \text{ m}^3} = 2.46 \times 10^{25} \text{ atoms/m}^3.$$

For the mass density, I must know the mass of 1 mol of helium atoms, and so I multiply the mass of one helium atom by the number of atoms in 1 mol:

$$m = (6.647 \times 10^{-27} \text{ kg/atom})(6.022 \times 10^{23} \text{ atoms/mol})$$

$$= 4.003 \times 10^{-3} \text{ kg/mol}.$$

Equation 1.4 then yields

$$\rho = \frac{4.003 \times 10^{-3} \text{ kg}}{24.5 \times 10^{-3} \text{ m}^3} = 0.163 \text{ kg/m}^3.$$

(*b*) For the liquid helium, the same reasoning gives me

$$n = \frac{6.022 \times 10^{23} \text{ atoms}}{32.0 \times 10^{-6} \text{ m}^3} = 1.88 \times 10^{28} \text{ atoms/m}^3$$

$$\rho = \frac{4.003 \times 10^{-3} \text{ kg}}{32.0 \times 10^{-6} \text{ m}^3} = 125 \text{ kg/m}^3.$$

Examples of non-SI units accepted for use along with SI units are the minute (1 min = 60 s), the hour (1 h = 3600 s), the liter (1 L = 10^{-3} m^3), and the metric ton (1 t = 10^3 kg).

A number of traditional, non-SI units are used in engineering; in various businesses, industries, sports, and trades; and in everyday life in the United States. Examples are inches, feet, yards, miles, acres, ounces, pints, gallons, and fluid ounces. These units are nondecimal, which makes it hard to interconvert them. When solving problems in this course, always begin by converting any quantities given in non-SI units to the SI equivalents. A conversion table is given in Appendix C.

The simplest way to convert from one unit to another is to write the conversion factor for the two units as a ratio. For example, in Appendix C, we see that 1 in. = 25.4 mm. By bringing the 25.4 mm to the left side of the equals sign, we can write this as either

$$\frac{1 \text{ in.}}{25.4 \text{ mm}} = 1 \quad \text{or} \quad \frac{25.4 \text{ mm}}{1 \text{ in.}} = 1. \tag{1.5}$$

Note that you *must* write the units in these expressions because without them you obtain the incorrect expressions $\frac{1}{25.4} = 1$ and $\frac{25.4}{1} = 1$. Because multiplying any number by 1 doesn't change the number, you can use these ratios to convert units. For example, to express 4.5 in. in millimeters, you multiply by the ratio on the right in Eq. 1.5 and cancel out the inches:

$$4.5 \text{ in.} = (4.5 \cancel{\text{in.}})\left(\frac{25.4 \text{ mm}}{1 \cancel{\text{in.}}}\right) = 4.5 \times 25.4 \text{ mm} = 1.1 \times 10^2 \text{ mm}. \qquad (1.6)$$

1.12 Why is the ratio on the left in Eq. 1.5 not suitable for converting inches to millimeters?

EXERCISE 1.7 Unit conversions

Convert each quantity to a quantity expressed either in meters or in meters raised to some power: (*a*) 4.5 in., (*b*) 3.2 acres, (*c*) 32 mi, (*d*) 3.0 pints.

SOLUTION I obtain my conversion factors from Appendix C.

(*a*) $$(4.5 \text{ in.})\left(\frac{2.54 \times 10^{-2} \text{ m}}{1 \text{ in.}}\right) = 1.1 \times 10^{-1} \text{ m}.$$

(*b*) $$(3.2 \text{ acres})\left(\frac{4.047 \times 10^3 \text{ m}^2}{1 \text{ acre}}\right) = 1.3 \times 10^4 \text{ m}^2.$$

(*c*) $$(32 \text{ mi})\left(\frac{1.609 \times 10^3 \text{ m}}{1 \text{ mi}}\right) = 5.1 \times 10^4 \text{ m}.$$

(*d*) $$(3.0 \text{ pints})\left(\frac{4.732 \times 10^{-4} \text{ m}^3}{1 \text{ pint}}\right) = 1.4 \times 10^{-3} \text{ m}^3.$$

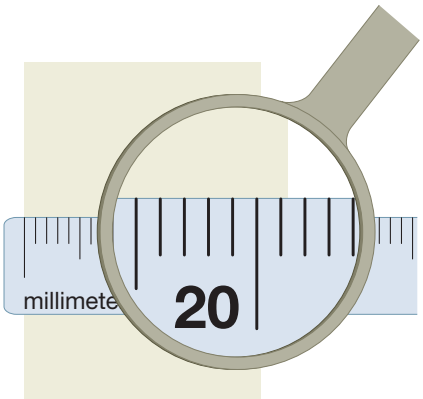
1.13 (*a*) Using what you know about the diameters of atoms from Section 1.3, estimate the length of one side of a cube made up of 1 mol of closely packed carbon atoms. (*b*) The mass density of graphite (a form of carbon) is 2.2×10^3 kg/m^3. By how much does the length you calculated in part *a* change when you do your calculation with this mass density value? Remember that 1 mol is approximately the number of atoms in 12×10^{-3} kg of carbon.

1.7 Significant digits

The numbers we deal with in physics fall into two categories: exact numbers that are known with complete certainty (integers, such as the 14 in "I have 14 books on my desk") and inexact numbers that result from measurements and are known to only within some finite precision. Consider, for example, using a ruler to measure the width of a piece of paper (**Figure 1.18**). The width falls between the ruler marks for 21 mm and 22 mm and is closer to 21 mm than 22 mm. We might guess that the width is about 21.3 mm, but we cannot be sure about the last digit without a better measurement method. By recording the width as 21 mm, we are indicating that the width actually lies between 20.5 mm and 21.5 mm. The value 21 mm is said to have two **significant digits**—digits that are known reliably.

By expressing a value with the proper number of significant digits, we can convey the precision to which that value is known. For numbers that don't contain any zeros, all digits shown are significant, which means that 21 has two significant digits, as just noted, and 21.3 has three significant digits (implying that the value lies between 21.25 and 21.35).

Figure 1.18 If you measure the width of a sticky note with the ruler shown, you can reliably read off two digits.



With numbers that contain zeros, the situation is more complicated. *Leading zeros,* which means any that come before the first nonzero digit, are never significant: 0.037 has two significant digits. Zeros that come between two nonzero digits, as in 0.602, are always significant. *Trailing zeros* are those that come after the last nonzero digit in a number, as in 3.500 and 20. Trailing zeros to the right of the decimal point are always significant: 25.10 has four significant digits. However, trailing zeros in numbers that do not contain a decimal point are ambiguous. For example, in the number 7900, it is not clear whether the trailing zeros are significant or not. The number of significant digits is at least two but could be three or four. To accurately convey the precision for such numbers, we use scientific notation, which lets us place significant zeros to the right of the decimal point: 7.900×10^3 has four significant digits, 7.90×10^3 has three, and 7.9×10^3 has two.

Important note: To simplify the notation in this book, we consider all trailing zeros in numbers that do not contain a decimal point to be significant: four significant digits for 3400, for instance, and two significant digits for 30.

Do not confuse significant digits with the number of decimal places or the number of digits: 0.000584 has three significant digits, six decimal places, and seven digits; 58.4 has three significant digits, one decimal place, and three digits. With the exception of fundamental constants such as the speed of light, most values in this book are given to two or three significant digits.

When you use a calculator for the math involved in solving physics problems, the calculator display usually shows many more digits than the number of significant digits allowed by the problem data. In such cases, you need to round your answer to the correct number of significant digits. If the digit just to the right of the last significant digit you are allowed is less than 5, report your last significant digit as it appears on the calculator display. If the digit just to the right of the last significant digit is 5 or greater, increase your last significant digit by 1. For example, if you are allowed two significant digits, 1.356 rounds to 1.4, 2.5199 rounds to 2.5, and 7.95 rounds to 8.0.

Multiple roundings can yield an accumulation of errors, and therefore it is best to wait until you have obtained the final result in a multistep calculation before rounding. In intermediate results, therefore, retain a few more digits than what your input quantities have and then round off to the correct number of significant digits only in your final result.

EXERCISE 1.8 Significant digits

(*a*) How many significant digits are there in 403.54 kg, 3.010×10^{57} m, 2.43×10^{-3} s, 14.00 μm, 0.0140 s, 5300 kg? (*b*) Round 12,300 kg and 0.0125 s to two significant digits.

SOLUTION (*a*) 403.54 kg has five, 3.010×10^{57} m has four, 2.43×10^{-3} s has three, 14.00 μm has four, 0.0140 s has three, 5300 kg has four in this book (but is considered ambiguous in general).

(*b*) 1.2×10^4 kg (or 12 Mg); 0.013 s (or 1.3×10^{-3} s or 1.3 ms).

Suppose you measure the mass and volume of some object to be $m = 1.2$ kg and $V = 0.123$ m^3, and you wish to use these values to compute the object's mass density. If you substitute the measured values into Eq. 1.4 and carry out the division on your calculator, you get

$$\rho \equiv \frac{m}{V} = \frac{1.2 \text{ kg}}{0.123 \text{ m}^3} = 9.756\,097\,56 \text{ kg/m}^3. \tag{1.7}$$

You measured the mass to two significant digits and the volume to three significant digits, but your calculator suggests you have determined the density to a precision of better than one part in a billion! The rigorous way to deal with uncertainties in measurements is to recalculate any computed value by using the high and low uncertainties for each value in the calculation, which is a time-consuming task. For this reason, we shall use two shortcuts. The first deals with multiplication and division:

When multiplying or dividing quantities, the number of significant digits in the result is the same as the number of significant digits in the input quantity that has the fewest significant digits.

In your calculation of the mass density in Eq. 1.7, the input quantity that has the fewest significant digits is the mass (two significant digits), and so the answer should be rounded to two significant digits:

$$\rho \equiv \frac{m}{V} = \frac{1.2 \text{ kg}}{0.123 \text{ m}^3} = 9.8 \text{ kg/m}^3. \tag{1.8}$$

Now suppose you determine the mass of two parts of an object and obtain 105 kg for one part and 0.01 kg for the other (three and one significant digit, respectively). Adding the two measured values yields a value containing five significant digits:

$$105 \text{ kg} + 0.01 \text{ kg} = 105.01 \text{ kg}. \tag{1.9}$$

Given that the precision in the value 105 kg is ± 0.5 kg, it makes no sense to report the final result to five significant digits. Rounding so that the number of significant digits in the sum is the same as the number of significant digits in the input value that has the fewest significant digits would yield 100 kg, which deviates from the "actual" value (around 105 kg) by significantly more than the precision in the individual measurements (± 0.5 kg and ± 0.005 kg). To solve this problem, whenever we add or subtract numerical values, we focus on the number of *decimal places*:

When adding or subtracting quantities, the number of decimal places in the result is the same as the number of decimal places in the input quantity that has the fewest decimal places.

For the addition in Eq. 1.9, the input quantity that has the fewest decimal places is 105 kg (zero decimal places), and so

$$105 \text{ kg} + 0.01 \text{ kg} = 105 \text{ kg}. \tag{1.10}$$

Bear in mind that *counting numbers* (that is, integers) are exact and so have an infinite number of significant digits. For example, the product of the width of two buttons each 15.7 mm wide is $2 \times 15.7 \text{ mm} = 31.4 \text{ mm}$ (not 3×10^1 mm because the 2 is a counting number).

EXERCISE 1.9 Significant digits and calculations

Calculate: (*a*) $f = a/(bc)$, where $a = 2.34 \text{ mm}^2$, $b = 54.26$ m, and $c = 0.14$ μm; (*b*) $g = kt^3$, where $k = 1.208 \times 10^{-2} \text{ s}^{-3}$ and $t = 2.84$ s; (*c*) $f + g$; (*d*) the sum of $b = 54.26$ m and $c = 1.4$ mm; (*e*) $h = k(m - n)$, where $k = 1.252$, $m = 32.21$, and $n = 32.1$.

(Continued)

SOLUTION (*a*) I first need to convert the millimeters-squared in $a = 2.34\ \text{mm}^2$ to meters-squared and the micrometers in $c = 0.14\ \mu\text{m}$ to meters:

$$2.34\ \text{mm}^2\left(\frac{1\ \text{m}}{1\times10^3\ \text{mm}}\right)^2 = 2.34\ \text{mm}^2\left(\frac{1\ \text{m}^2}{1\times10^6\ \text{mm}^2}\right) = 2.34\times10^{-6}\ \text{m}^2$$

$$0.14\ \mu\text{m}\left(\frac{1\ \text{m}}{1\times10^6\ \mu\text{m}}\right) = 0.14\times10^{-6}\ \text{m}.$$

Substituting the values given, I get

$$f = \frac{2.34\times10^{-6}\ \text{m}^2}{(54.26\ \text{m})\,(0.14\times10^{-6}\ \text{m})} = 0.30804,$$

which I must round to two significant digits, 0.31, because that is the number of significant digits in the value given for *c*.

(*b*) $g = (1.208\times10^{-2}\ \text{s}^{-3})(2.84\ \text{s})^3 = 0.2767$, which I must round to three significant digits, 0.277, because of *t*.

(*c*) $f + g = 0.30804 + 0.2767 = 0.58474$, which I round to 0.58 because I must report only two decimal places, limited by $c = 0.14\ \mu\text{m}$. Note that if I had added the rounded values for *f* and *g*, my result reported to two significant digits would not be 0.58: $f + g = 0.31 + 0.277 = 0.587$, which rounds to 0.59. Be careful—reporting the correct number of significant digits can be tricky.

(*d*) $b + c = 54.26\ \text{m} + 0.0014\ \text{m} = 54.26\ \text{m}$ (reported to two decimal places, limited by the 54.26 value).

(*e*) $h = 1.252(32.21 - 32.1) = 0.1$ (reported to one significant digit, limited by the 32.1 value).

The number zero requires special consideration. When some physical quantity (other than temperature) is exactly zero, we denote this quantity by a zero *without units*. The speed of an object that is not moving is exactly zero and is therefore denoted by $v = 0$ (no units). However, if we measure a speed as zero to two significant digits, we write $v = 0.0\ \text{m/s}$ (note the units), implying that the value is zero to within 0.05 m/s.



1.14 (*a*) Express the circumference of a circle of radius $R = 27.3$ mm with the correct number of significant digits. (*b*) Let $a = 12.3$, $b = 3.241$, and $c = 55.74$. Compute $a + b + c$. (*c*) Let $m = 4.00$, $n = 3.00$, and $k = 7$ (exact). Compute $f = m^2/k$, $g = n^2/k$, and $f + g$.

1.8 Solving problems

You encounter problems every day. You are in a rush to get somewhere but can't find your car keys. You run out of flour while baking a birthday cake, and the supermarket is closed. Your flight is canceled on your way to a job interview. You want to buy those great new shoes, but there's no money in the bank. There is no formula for solving these or any other types of problems (if there were a formula, you wouldn't have a problem!). Because real-world problems defy one-shot solutions, we must break them down into smaller parts and solve them in steps. In this section we develop a four-step problem-solving strategy that will help you tackle a broad variety of problems (the procedure is summarized in the box "Solving problems" on page 41).

Most physics problems are formulated in words. This is true not only of the problems in this book but also of questions you might have about the world around you. Standing on the rim of the Grand Canyon, you might wonder, If I drop a stone, how long does it take to hit the bottom?, or while watching the Olympics, you might wonder, Is there a physical limit to how high an athlete can jump?

❶ GETTING STARTED

Because it's not clear which road leads you most efficiently to the answer to a given problem, the first step of our problem-solving strategy, *getting started,* is the most difficult one. It is therefore useful to begin with something you *can* do: Organize the information given and be sure you are clear on exactly what is asked for in the problem. To clarify the goal of the problem in your mind, it is crucial to begin by visualizing the situation: Sketch, describe in your own words, list, and/or tabulate the main features of the problem to relieve yourself from having to hold all this information in memory. In making a sketch, be sure to follow the guidelines given in Section 1.5: Make your rendering as simple as possible and show all relevant numerical information in the sketch. Recasting your problem visually or verbally forces you to spell out what you want to accomplish and often automatically leads to ideas on how to solve the problem.

Once you have the information organized, ask yourself which concepts and principles apply. Throughout this book, we develop principle-specific procedures for solving problems. By determining which principles apply to a problem you are working on, you can identify which problem-solving procedures apply to your problem.

Finally, you must determine whether or not you have all the information necessary to solve the problem. Sometimes you may have to supplement the information provided with things you know about the situation at hand. Furthermore, because no problem can be formulated with absolute precision, you will often need to make a number of simplifying assumptions. This lack of precision is often a source of frustration for anyone beginning to solve physics problems. How do you know which assumptions are valid for a given problem? Do not let this question trouble you—as long as you are aware of the assumptions you make, you can always reexamine them once you have obtained an answer and then refine the assumptions and the solution.

❷ DEVISE PLAN

Next you must *devise a plan* for solving your problem, which means spelling out what you must do to solve the problem. Are there any physical relationships or equations that you can apply to determine the information you are trying to obtain? A good plan is to outline the steps you need to take to obtain a solution. For some (but certainly not all) problems, these steps are carried out mathematically.

❸ EXECUTE PLAN

In the third step, you *execute your plan* by following the steps you have outlined, substituting the information given and carrying out any mathematical operations necessary to isolate the quantity you wish to determine. In problems involving numerical values, you should solve for an algebraic answer first, waiting until the final step to substitute numerical values and obtain a numerical answer. The only exception to this algebraic-answer-first approach is that, in order to simplify your algebraic expressions, you should eliminate any quantities that are zero from them as soon as possible.

In the final step, substituting numerical values into your algebraic expression, do not forget to show the correct units on all numerical values and to carry these units through the calculation.

Once you have obtained an answer, you should check your work for these five important points (which you can remember via the acronym *VENUS*):

1. **V**ectors/scalars used correctly? As you will see in Chapter 2, the quantities we deal with in this book fall into two classes, vectors and scalars. Make sure you have expressed your answer in the appropriate form. (In this chapter all quantities are scalars.)
2. **E**very question asked in problem statement answered? Reread the problem statement to make sure you have completely answered every question.
3. **N**o unknown quantities in answers? If your answer is an algebraic expression, be sure that every variable appearing in terms to the right of the equals sign has been given as a known quantity in the problem statement. Example: "A car moving at speed v . . . " means that v in this problem is a known quantity.
4. **U**nits correct? Each numerical answer should be expressed as a number with the appropriate units for that quantity (Section 1.6). When your answer is an algebraic expression, be sure that the combined units of all terms to the right of the equals sign work out to be equivalent to the unit of the quantity to the left of the equals sign.
5. **S**ignificant digits justified? The number of significant digits reported in your answer should reflect the number of significant digits of the quantities given in the problem statement (Section 1.7).

As a reminder to yourself, put a checkmark beside each answer to indicate that you checked these five points.

❹ EVALUATE RESULT

You may think you are done, but there is one final—and very important—step left: *Evaluate your answer*. In this step, you should reflect on your work, examine the result, and determine whether your answer is reasonable (or, if you can't tell whether or not it is reasonable, at least make sure that it is not unreasonable). The first thing you might try is seeing whether your answer conforms to what you expect based on your situation sketch, diagram, or other information given. If any part of your answer is unexpected, go back and check your math and any assumptions you made.

If your result is an algebraic expression or if your numerical answer comes directly from an algebraic expression, check whether or not that expression behaves as you expect with changes in physical quantities. For example, say you obtain the expression $v = k/t$ for the speed of an object and you know from experience that for this object the speed is zero for large values of t. If you let t go to infinity, your expression should indeed give the right answer: $v = k/\infty = 0$. If it didn't, you must go back and check your math and any assumptions you made.

Some problems can be solved in more than one way. If you solve your problem in two ways and obtain the same answer, you can be pretty sure your answer is correct. If you don't get the same answer from the two ways, you'd better go back and check your work!

If none of the evaluation methods can be applied to your answer, you should at least verify that there is nothing obviously wrong with it by checking that the algebraic signs you obtain make sense and that any numerical value has the correct order of magnitude. For example, if the algebraic sign of your solution tells you that an object moves to the left when you know it should move to the right, you know your answer can't be right. You can find some techniques for estimating orders of magnitude

PROCEDURE: Solving problems

Although there is no set approach when solving problems, it helps to break things down into several steps whenever you are working a physics problem. Throughout this book, we use the four-step procedure summarized here to solve problems. For a more detailed description of each step, see Section 1.8.

1. **Getting started.** Begin by carefully analyzing the information given and determining in your own words what question or task is being asked of you. Organize the information by making a sketch of the situation or putting data in tabular form. Determine which physics concepts apply, and note any assumptions you are making.
2. **Devise plan.** Decide what you must do to solve the problem. First determine which physical relationships or equations you need, and then determine the order in which you will use them. Make sure you have a sufficient number of equations to solve for all unknowns.
3. **Execute plan.** Execute your plan, and then check your work for the following five important points:

   Vectors/scalars used correctly?
   Every question asked in problem statement answered?
   No unknown quantities in answers?
   Units correct?
   Significant digits justified?

   As a reminder to yourself, put a checkmark beside each answer to indicate that you checked these five points.
4. **Evaluate result.** There are several ways to check whether an answer is reasonable. One way is to make sure your answer conforms to what you expect based on your sketch and the information given. If your answer is an algebraic expression, check to be sure the expression gives the correct trend or answer for special (limiting) cases for which you already know the answer. Sometimes there may be an alternative approach to solving the problem; if so, use it to see whether or not you get the same answer. If any of these tests yields an unexpected result, go back and check your math and any assumptions you made. If none of these checks can be applied to your problem, check the algebraic signs and order of magnitude.

in the next section. Using these techniques, you should be able to obtain reasonable upper and lower bounds for your answer. Provided your answer does not lie outside this range, you know there is not something obviously wrong with it.

With the exception of one-step exercises that are applications of procedures, the examples in this book follow this four-step procedure.

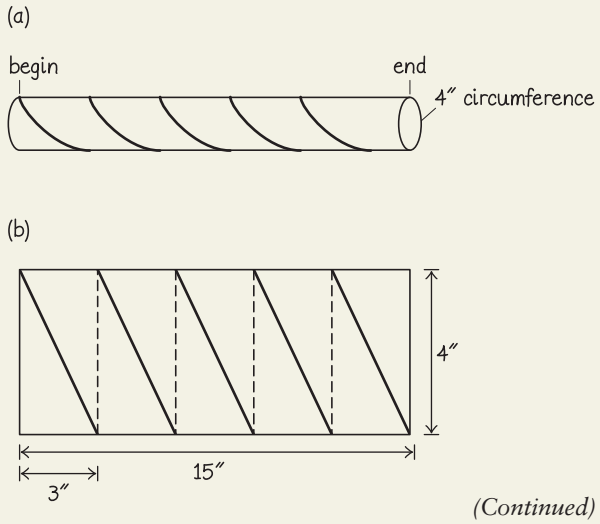
EXAMPLE 1.10 Length of string

A certain cylinder is 15 in. long and has a circumference of 4.0 in. When one end of a piece of string is attached to the left end of the cylinder and the string is then wound exactly five times around, the string just reaches to the cylinder's right end. What is the length of the string?

❶ **GETTING STARTED** I begin by making a sketch of the situation (**Figure 1.19*a***). The shortest length is going to be obtained when the string is wound uniformly, with each of the five windings running one-fifth of the length of the cylinder. I need to find a way to determine the length of this spiraling string. This appears to be a geometry problem, but the surface of the cylinder is curved, not flat.

❷ **DEVISE PLAN** If I imagine drawing a straight line between the endpoints of the string, cutting the cylinder surface along that line, and then flattening the cylinder out, I obtain the situation shown in Figure 1.19*b*. Now the solution is clear: The total length of the string is the sum of the hypotenuses of five identical triangles.

Figure 1.19



(Continued)

❸ **EXECUTE PLAN** The width of each triangle is equal to one-fifth of the length of the cylinder: $(15 \text{ in.})/5 = 3.0 \text{ in.}$ The height of each triangle is equal to the circumference of the cylinder (4.0 in.). With these values, I can now use the Pythagorean theorem to determine the length ℓ of the hypotenuse of each triangle:

$$\ell = \sqrt{(3.0 \text{ in.})^2 + (4.0 \text{ in.})^2} = 5.0 \text{ in.}$$

The length of the string is equal to five times this value, or 25 in. ✔*

❹ **EVALUATE RESULT** The answer I obtain should be larger than the string length I would obtain if the string were wound around the cylinder in one place, which would give a string length of five times the circumference: $5 \times 4.0 \text{ in.} = 20 \text{ in.}$ My answer should be smaller than what I would get if the string were wound five times around the cylinder at one end and then run along the cylinder length to the other end, which would give a string length of five times the circumference plus the length of the cylinder: $(5 \times 4.0 \text{ in.}) + 15 \text{ in.} = 35 \text{ in.}$ The answer I obtained does indeed lie between these values, giving me confidence that it is correct.

*The checkmark shows that I checked my answer regarding the five important details listed on page 40: Vector/scalar: all quantities scalar; Every question answered: yes; No unknown quantities in answers: no; Units correct: yes, inches; Significant digits: everything given to two significant digits.

1.9 Developing a feel

Making order-of-magnitude estimates is possibly the most important skill to retain from a physics course because this is a skill you will use long after you leave your last classroom, no matter what you plan for a career. Economists, for instance, carry out order-of-magnitude estimates to predict market trends, and business owners might carry out such estimates to validate a business plan, to name just two possibilities.

Here in this physics course, to get back to the present, order-of-magnitude estimates allow you to develop a feel for a problem before you try to solve it. They also allow you to validate your answer: If your order-of-magnitude estimate and your answer disagree, you know to go back and reevaluate your work.

Order-of-magnitude estimates involve determining relationships between known and unknown quantities, making gross oversimplifications, combining assumptions and ideas, and even outright guessing. In general, there are multiple ways to obtain such estimates. The uncertainty in how to carry out an estimate is intimidating at first, and the lack of accuracy may appear unscientific to you. It isn't—estimation is an important aspect of science, and it is a great way to obtain physical insight and at the same time train your analytical and outside-the-box thinking skills.

For example, suppose we need to determine the number of cells in the human body. That we can't see individual cells tells us they must be very small, and therefore the number is probably very large. We could *guess* the number of cells, but the answer would probably be meaningless because we have no way of assessing its accuracy without looking it up somewhere. A billion? Ten billion? Hundred billion? What do you think?

To obtain an order-of-magnitude estimate, the first thing to do is to think of a strategy. For the number of cells in the human body, for example, if we knew the volume of a typical human body and the volume of a typical cell, we could divide the former by the latter to obtain our answer. Knowing neither volume, however, we need to estimate both.

There are several ways to estimate the volume of the human body. One way is to model the body as a cylinder (**Figure 1.20**). For the height of the cylinder, we take the height of a typical human, say $h_{\text{human}} \approx 1.7$ m. The symbol $\approx$ indicates that this value is an approximation. For the diameter, we take $d_{\text{human}} \approx 0.2$ m. That makes the diameter less than the left-to-right thickness of the body but greater than the front-to-back thickness, which should be in the right ballpark. This gives a volume of

$$V_{\text{human}} = \left(\frac{1}{4}\pi d_{\text{human}}^2\right) h_{\text{human}} \approx 0.05 \text{ m}^3.$$

Figure 1.20 To determine the volume of the human body, we approximate the body as a cylinder.



QUANTITATIVE TOOLS

Another way is to realize that the human body has roughly the same mass density as water—about 1 kg/L, which is

$$1 \text{ kg/L} \left(\frac{1 \text{ L}}{1 \times 10^{-3} \text{ m}^3} \right) = 1000 \text{ kg/m}^3.$$

The average human has a mass of about 70 kg, which corresponds to 70 L of water, and thus $V_{\text{human}} \approx (70 \text{ L})(1 \times 10^{-3} \text{ m}^3/1 \text{ L}) = 0.07 \text{ m}^3$, which is in the same ballpark as the estimate we got using the body-as-cylinder approach.

You may feel that there is too big a difference between 0.05 m^3 and 0.07 m^3. Indeed, if you were shopping for clothing, you'd want your clothing to fit your body better than that. However, here we are interested in only the order of magnitude. Either estimate is good enough for this purpose because for both the order of magnitude is the same: $0.05 \text{ m}^3 = 5 \times 10^{-2} \text{ m}^3$, which rounds to $10 \times 10^{-2} \text{ m}^3$, and so the order of magnitude is 10^{-1} m^3; $0.07 \text{ m}^3 = 7 \times 10^{-2} \text{ m}^3$, which also rounds to $10 \times 10^{-2} \text{ m}^3$, giving the same order of magnitude: 10^{-1} m^3.

Next we need to estimate the volume of a typical human cell. You may remember from biology class that the diameter of a typical cell is around $20 \ \mu\text{m} = 20 \times 10^{-6}$ m. Or, if you never learned cell sizes, you may recall that you can clearly see cells under a microscope and that a microscope can see objects as small as 10^{-6} m. To see the structure of a cell clearly under a microscope, the cell would need to be, say, at least 50 times the minimum size the microscope can handle, making our cell-diameter estimate 50×10^{-6} m. Although you may think there is a big difference between 20×10^{-6} m and 50×10^{-6} m, these two numbers are only a factor of 2.5 apart, which is much less than a factor of 10 ("one order of magnitude").

Cells viewed under a microscope tend to spread out on the glass and are thus less high than they are wide. Let's for simplicity assume they are disk-shaped and that their height is about one-quarter of their diameter. That means they have a diameter of 20×10^{-6} m and a height of 5×10^{-6} m, and so $V_{\text{cell}} = \left(\frac{1}{4}\pi d_{\text{cell}}^2\right) h_{\text{cell}} \approx 1.6 \times 10^{-15} \text{ m}^3$. The number of cells is thus $V_{\text{human}}/V_{\text{cell}} \approx (0.05 \text{ m}^3)/(1.6 \times 10^{-15} \text{ m}^3) \approx 3.1 \times 10^{13}$. Given the roughness of the estimates and assumptions, I give the final answer as an order of magnitude: 10^{14} (about 100 trillion cells).

I can be quite confident that the number of cells is somewhere between 10^{13} (10 trillion) and 10^{15} (1000 trillion), both substantially greater than the largest guess I offered earlier (100 billion = 10^{11}). The difference between that largest guess and the order-of-magnitude estimate is that I have no idea how much confidence to place in the guess, whereas I can feel reasonably certain about the result of my order-of-magnitude estimate.

You may think that this exercise is outrageous and that it is impossible to have faith in any order-of-magnitude value because it requires so many assumptions and estimates. However, the fact that we combine so many assumptions and estimates is precisely why the method works: We are unlikely to be either too high on every estimate or too low on every estimate. Instead, we unwittingly overestimate some values and underestimate others, and the over- and underestimates tend to cancel out. The result is that the final answer very likely is in the right ballpark. What is more important is that, in the process of obtaining an order-of-magnitude value, we develop insights and some feeling for the quantities involved.

Figure 1.21 shows in graphical form the process by which we obtained our estimate of the number of cells N_{cell} in the human body. We got to our answer through a combination of *devising a strategy* (using a ratio of volumes to determine the number of cells), *simplifying* (treating the human body as a cylinder and the cell as a disk), *estimating* (the diameter of the human body),

and using snippets of *knowledge*. Figure 1.21 also shows why this approach can be intimidating at first: To determine a single target unknown, N_{cell}, we have to divide the problem into a larger number of questions, an approach that makes the number of unknowns increase initially! The point of this dividing, of course, is to arrive at values we know (or can guess).

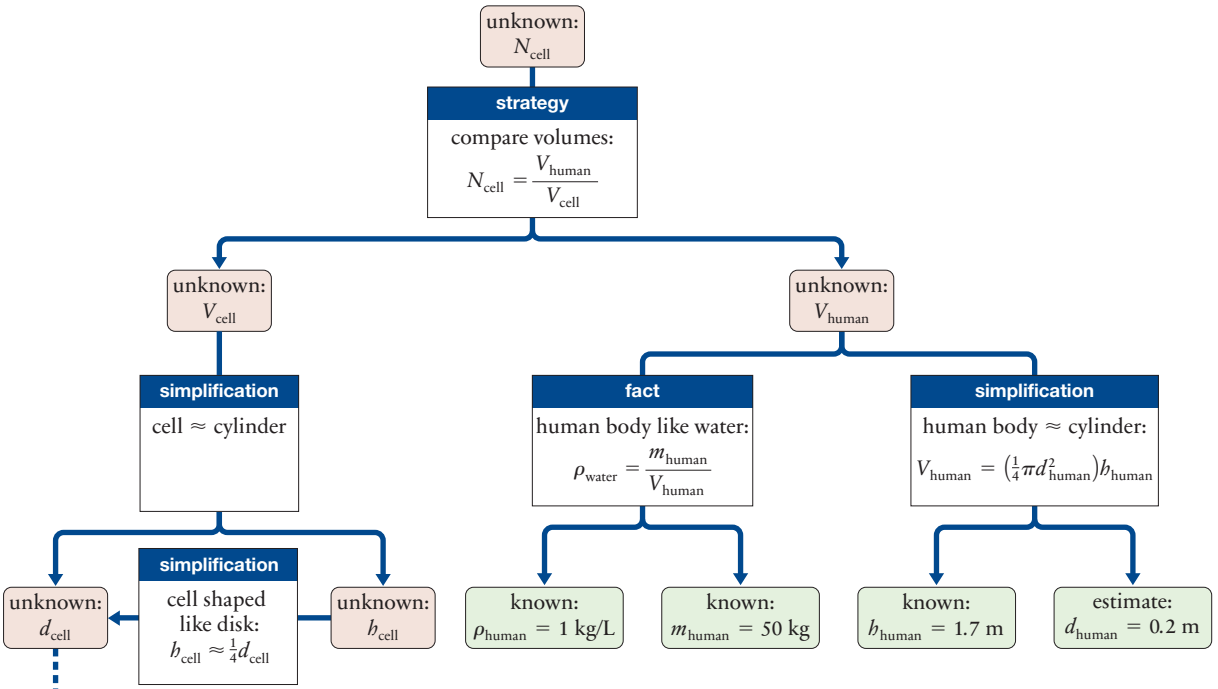
The diagram in Figure 1.21*a* reduces the problem of obtaining N_{cell} to one of obtaining another value, d_{cell}. If you know the value of d_{cell} from your biology course, you are done. If you don't know the value, you can *translate* the problem into another one, as illustrated in Figure 1.21*b*. If you get stuck, guess upper and lower bounds for your unknowns, use the middle value to obtain an answer, and then reevaluate to see how your *guesstimate* affects the final answer.

Making order-of-magnitude estimates requires practice. As you go through the exercise more often, you'll gain confidence, and the process will come more naturally. An important point is not to worry too much about accuracy when working out orders of magnitude: A factor of 10 doesn't matter. You'll be surprised how far you can get with this technique, some resourcefulness, and commonly available knowledge.

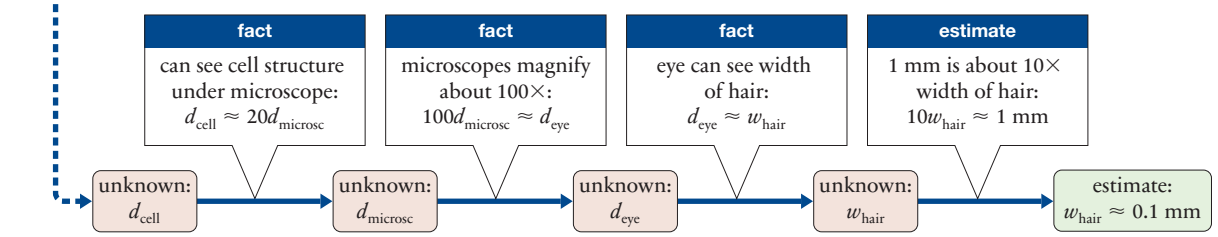
1.15 To work an order-of-magnitude problem, you need the value of the mass density of water, but you don't know what that value is. Design either a translation strategy (Figure 1.21*b*) or a division strategy (Figure 1.21*a*) for making an order-of-magnitude estimate of this mass density.

Figure 1.21 The process by which we estimated the number of cells in the human body.

(*a*) A division strategy



(*b*) A translation strategy



EXERCISE 1.11 Selling books

Consider writing a book for high-school graduates in the United States who are planning to take a year off—sometimes called a gap year—before enrolling in college. There are currently no similar books. Estimate how many such books might be purchased annually.

SOLUTION I can estimate the number of books sold by estimating the percentage of college-bound high-school graduates who take a gap year. I know two things: (1) Not all graduates do this, and (2) there is a rising interest in doing it. I estimate that 5% of high-school graduates take a gap year before going to college, but I still have to estimate the number of college-bound high-school students in the United States. I know that approximately 50% of high-school graduates go on to college, but how many high-school graduates are there each year?

If I assume every person in the country graduates from high school when she or he is 18, I've reduced this part of the problem to determining the number of 18-year-olds in the country. I know that the total population is 300 million and that the average life span is about 75 years. If I assume that age is uniformly distributed, then 1/75 of the U.S. population should be 18 years old at any given instant (this is an oversimplification but is good enough for an order-of-magnitude estimate). So the number of high-school graduates in a given year is $(1/75)(3 \times 10^8) = 4 \times 10^6$, and 5% of this is 2×10^5. Thus the number of students taking a gap year each year is approximately 200,000, which strikes me as a reasonable number. If 5% of these students buy the proposed book, about 10,000 copies will be purchased each year.



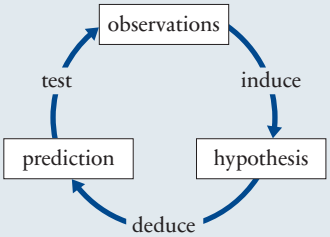
1.16 Make an order-of-magnitude estimate of the mass of Earth in kilograms.

Chapter Summary

The scientific method (Section 1.1)

Concepts The **scientific method** is an iterative process for going from observations to a hypothesis to an experimentally validated theory. If the predictions made by a hypothesis prove accurate after repeated experimental tests, the hypothesis is called a **theory** or a **law,** but it always remains subject to additional experimental testing.

Quantitative Tools



Symmetry (Section 1.2)

Concepts An object exhibits **symmetry** when certain operations can be performed on it without changing its appearance. Important examples are *translational symmetry* (movement from one location to another), *rotational symmetry* (rotation about a fixed axis), and *reflection symmetry* (reflection in a mirror). The concept of symmetry applies both to objects and to physical laws.

Some basic physical quantities and their units (Sections 1.3, 1.4, 1.6)

Concepts **Length** is a distance or extent in space. The SI (International System) base unit of length is the **meter** (m).

Time is a property that allows us to determine the sequence in which related events occur. The SI base unit of time is the **second** (s). The **principle of causality** says that whenever event A causes an event B, all observers see event A happen before event B.

Density is a measure of how much of some substance there is in a given volume.

Quantitative Tools If there are N objects in a volume V, then the *number density* n of these objects is

$$n \equiv \frac{N}{V}. \tag{1.3}$$

If an object of mass m occupies a volume V, then the *mass density* ρ of this object is

$$\rho \equiv \frac{m}{V}. \tag{1.4}$$

To convert one unit to an equivalent unit, multiply the quantity whose unit you want to convert by one or more appropriate *conversion factors*. Each conversion factor must equal one, and any combination of conversion factors used must cancel the original unit and replace it with the desired unit. For example, converting 2.0 hours to seconds, we have

$$2.0 \text{ h} \times \frac{60 \text{ min}}{1 \text{ h}} \times \frac{60 \text{ s}}{1 \text{ min}} = 7.2 \times 10^3 \text{ s}.$$

Representations (Section 1.5)

Concepts Physicists use many types of representations in making models and solving problems. Rough sketches and detailed diagrams are generally useful, and often crucial, to this process. Graphs are useful for visualizing relationships between physical quantities. Mathematical expressions represent models and problems concisely and permit the use of mathematical techniques.

Quantitative Tools As you construct a model, begin with a simple visual representation (example: represent a cow with a dot) and add details as needed to represent additional features that prove important.

Significant digits (Section 1.7)

Concepts The **significant digits** in a number are the digits that are reliably known.

Quantitative Tools If a number contains no zeros, then all the digits shown are significant: 345 has three significant digits; 6783 has four significant digits.

For numbers that contain zeros:

- Zeros between two nonzero digits are significant: 4.03 has three significant digits.
- Trailing digits to the right of the decimal point are significant: 4.9000 has five significant digits.
- Leading zeros before the first nonzero digit are not significant: 0.000175 has three significant digits.
- In this book, trailing zeros in numbers without a decimal point are significant: 8500 has four significant digits.

The number of significant digits in a product or quotient is the same as the number of significant digits in the input quantity that has the *fewest significant digits:* $0.10 \times 3.215 = 0.32$.

The number of decimal places in a sum or difference is the same as the number of decimal places in the input quantity that has the *fewest decimal places:* $3.1 + 0.32 = 3.4$.

Solving problems and developing a feel (Sections 1.8, 1.9)

Concepts Strategy for solving problems:

1. **Getting started.** Analyze and organize the information and determine what is being asked of you. A sketch or table is often helpful. Decide which physics concepts apply.
2. **Devise plan.** Determine the physical relationships and equations necessary to solve the problem. Then outline the steps you think will lead to the solution.
3. **Execute plan.** Carry out the calculations, and then check your work using the following points:
   - Vectors/scalars used correctly?
   - Every question answered?
   - No unknown quantities in answers?
   - Units correct?
   - Significant digits justified?
4. **Evaluate result.** Determine whether or not the answer is reasonable.

To develop a feel for the approximate size of a calculated quantity, make an **order-of-magnitude** estimate, which means a calculation rounded to the nearest power of ten.

Quantitative Tools Determining order of magnitude

Example 1:

4200 is 4.200×10^3.

Round the coefficient 4.200 to 10 (because it is greater than 3), so that 4.200×10^3 becomes $10 \times 10^3 = 1 \times 10^4$.

The order of magnitude is 10^4.

Example 2:

0.027 is 2.7×10^{-2}.

Round the coefficient 2.7 to 1 (because it is less than 3), so that 2.7×10^{-2} becomes 1×10^{-2}.

The order of magnitude is 10^{-2}.

Strategy to compute order-of-magnitude estimates:

- Simplify the problem.
- Break it into smaller parts that are easier to estimate.
- Build your estimate from quantities that you know or can easily obtain.

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Solving context-rich problems

The problems labeled CR are *context-rich* problems—problems that are more like those in the everyday world. These questions are typically embedded in a short narrative and often do not specify what variables you must calculate to answer the question. So, rather than asking What is the mass of . . . ? (telling you that you must calculate the mass of some object), a context-rich problem may ask Do you accept the bet? (leaving it to you to determine what to calculate). The problem typically contains extraneous information, and you may need to supply some missing information, either by estimation or by looking up values.

Like all problems, context-rich problems should be solved using the four-step procedure outlined in the Procedure box in this chapter. Because context-rich problem statements are never broken down into parts, the first two steps ("Getting started" and "Devise plan") are particularly important.

Context-rich problems will sharpen the skills you need to solve everyday problems, where you need to move toward a goal, but the path you must take is not immediately clear (if it were clear, there wouldn't be a problem). The information available to you may be a bit sketchy or contradictory, there may be more than one way to approach a situation, and some ways may be more fruitful than others, but that may not be clear at the outset. So consider context-rich problems as an opportunity to stretch your problem-solving abilities.

Dots indicate difficulty level of problems: • *= easy,* •• *=* ***intermediate,*** ••• *= hard;* CR *= context-rich problem.*

1.1 The scientific method

1. In a discussion of the technological capabilities of ancient civilizations, a friend offers this hypothesis: "Some ancient civilizations built sophisticated instruments based on a mysterious type of radiation, but all of them have been lost." Which part of his statement is most likely to keep the statement from satisfying the criteria for a scientific hypothesis? •
2. An advertisement for a food product states that it contains 50% less fat per serving than a competing product. What assumptions are you making if you accept the claim as valid? •
3. You are asked to predict the next item in the following sequence of integers: 1, 2, 3. If your prediction is 4, what assumptions did you make? ••
4. Checkpoint 1.1 states that two coins together have a value of 30 cents but one of them is not a nickel. If the checkpoint said instead that "neither coin is worth 5 cents," what hidden assumptions might prevent you from obtaining a solution? ••
5. In a 4-by-4 Sudoku puzzle with one 2-by-2 subsquare filled in as shown in Figure P1.5, how many ways are there to complete the puzzle? (In Sudoku, the digits 1, 2, 3, 4 must appear only once in each row and once in each column of the 4-by-4 square and only once in each subsquare.) •••

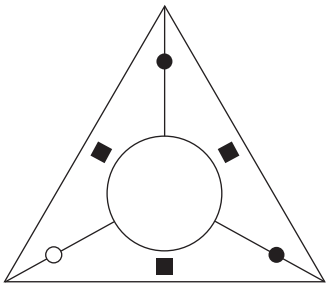
Figure P1.5

4	3		
1	2		

1.2 Symmetry

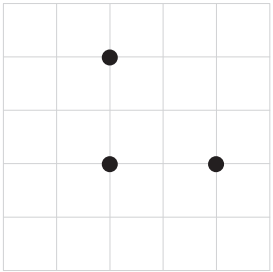
6. How many axes of reflection symmetry does the triangle in Figure P1.6 have? Restrict your answer to axes that lie in the plane of the triangle. •

Figure P1.6



7. How many axes of rotational symmetry does a right circular cone have? •
8. Three identical coins are positioned on a grid as shown in Figure P1.8. Where would you place a fourth coin to form a coin arrangement that has both reflection symmetry and a 90° rotational symmetry? •

Figure P1.8



9. Describe the reflection symmetry of each of the letters in the word below: ••

T A B L E S

10. Which types of geometrical symmetry mentioned in this chapter does a sphere have? ••
11. How many axes of reflection symmetry does a cube have? How many axes of rotational symmetry? ••
12. Suppose you have a disk of cardboard with two perpendicular diameters dividing the disk into four quadrants. The first and third quadrants taken clockwise from the topmost point of the disk are red, whereas the other two are white. How many axes of reflection does the disk have? Restrict your answer to axes in the plane of the disk. Does the disk have rotational symmetry about any axis perpendicular to the plane of the disk? ••
13. A cube is placed on a table with a green side facing up. You know that the side facing down is also green but has a red dot in the middle of the green face. From where you are sitting, you can see that two adjacent vertical sides are one red and the other blue. What is the maximum number of axes of reflection symmetry possible? •••

1.3 Matter and the universe

14. Traveling at 299,792,458 m/s, how far does light travel in 78 years? •
15. Light takes approximately 500 seconds to travel from the Sun to the Earth. (*a*) What is this distance in kilometers? (*b*) How many Suns could fit side by side in this distance? Take the radius of the Sun to be 696,000 km. •
16. Figure 1.9 tells us that a human contains 10^{29} atoms and a blue whale contains 10^{32} atoms. Use these values to determine the ratio of the length of a blue whale to the height of a human. ••
17. A gastrotrich is a small aquatic organism with a life span of about 3 days, and a giant tortoise has an average life span of 100 years. Make an order-of-magnitude estimate of the number of gastrotrich lifetimes that are equal to one tortoise lifetime. In your calculation, use an order-of-magnitude value for the number of days in 1 year. ••
18. You measure the diameter of a drop of water to be 3 mm. Make an order-of-magnitude estimate of how many such drops are contained in all the water in your body. ••
19. If you could stack copies of your physics textbook until they reached the Moon, what is the order of magnitude of the number of books you would need? ••
20. How many water molecules are there in an Olympic swimming pool that is 50 m long, 25 m wide and 3.0 m deep? ••
21. Cube 1 has side ℓ_1 and volume V_1. Cube 2 has side $\ell_2 = 2\ell_1$. (*a*) When making order-of-magnitude estimates, by how many orders of magnitude is the volume of the second cube greater than the volume of the first cube? (*b*) Does the answer depend on the numerical value of ℓ_1? ••
22. Earth requires approximately 365 days to orbit the Sun along a path that can be approximated by a circle with a radius of 1.50×10^8 km. Make an order-of-magnitude estimate of the length of time it would take light to move once around this orbit. ••
23. While admiring a tree, you notice that the leaves form a more or less continuous spherical shell, with relatively few leaves growing in the shell interior. You estimate that the diameter of this shell is about 30 m. Each leaf is about 5 in. long and 3 in. wide. Make an order-of-magnitude estimate of the number of leaves on this tree. •••

1.4 Time and change

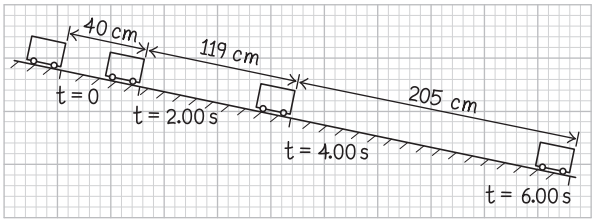
24. A human heart beats 72 times every minute. How many times, on average, does a heart beat during a healthy 70-year lifespan? (Ignore the fact that the rate at which the heart beats can vary widely with the level of physical activity.) •
25. You hear a peal of thunder and go to the window. You then see a flash of lightning. Is it reasonable to assume that the thunder caused the lightning? •
26. A second is defined as the duration of 9.19×10^9 periods of the radiation emitted by a cesium atom. What is the duration in seconds of a single period for this atom? •
27. While watching a railroad crossing, you observe that a crossing barrier is lowered about 30 s before a train passes. This happens every time for a long while, until suddenly after many times the barrier is not lowered but a train passes anyway! Discuss the causal relationship between the passing of a train and the lowering of the barrier. Does the single unusual occurrence make a difference to your answer? ••
28. Assume the radiation from the cesium atom mentioned in Problem 26 moves at the speed of light. How many meters does the radiation travel in the length of time corresponding to the period calculated in Problem 26? ••

1.5 Representations

29. Translate this statement into a mathematical expression, using any symbols you wish, but include a "key" to the meaning of each symbol: The wavelength of any electromagnetic radiation traveling in a vacuum is equal to the speed of light divided by the frequency of that radiation. •
30. A planar shape is formed by joining the ends of four straight-line segments, two with length ℓ and two with length 2ℓ. No loose ends are allowed, but crossings are allowed, and only two segment ends may join at any single point. The ends of identical-length segments cannot join each other. What shapes are possible? ••
31. Four straight-line segments of length ℓ lie in a plane and do not cross. One end of segment 1 is connected to one end of segment 2 such that they form a 30° interior angle. The other end of segment 2 is connected to one end of segment 3, also forming a 30° interior angle. Finally, the other end of segment 3 is connected to one end of segment 4, forming another 30° interior angle. What is the distance between the unconnected ends of segments 1 and 4? ••
32. You are arranging family members for a photograph, and you want them to stand in order of increasing height. Your uncle is half a foot shorter than your aunt, who is taller than your cousin. Your grandmother is 2 in. shorter than your grandfather. Your brother is 1 in. taller than your aunt and 3 in. taller than your cousin. Your grandfather is 1 cm taller than your aunt. List the relatives in order of increasing height. ••
33. The Jupiter–Sun distance is 778 million kilometers, and the Earth–Sun distance is 150 million kilometers. Suppose an imaginary line from Jupiter to the Sun forms a right angle with an imaginary line from Earth to the Sun. How long would it take light to travel from the Sun to a spaceship halfway along the straight-line path between Jupiter and Earth? ••

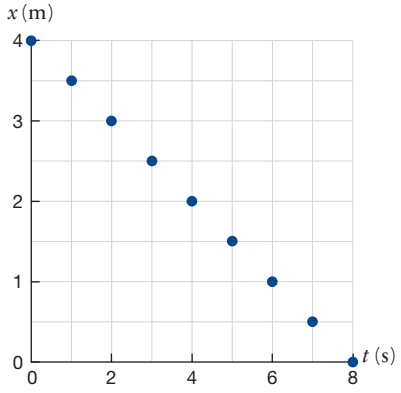
34. During a physics lab, you and your partner keep track of the position of a small cart while it is moving along a slope. The lab instructions ask you to prepare (*a*) a pictorial representation of the experiment, (*b*) a table recording the position of the cart measured every 2.00 seconds from its initial position, and (*c*) a graph that shows the position of the cart on the vertical axis and time on the horizontal axis. You produce the sketch shown in Figure P1.34. What should your table and your graph look like? ●●

Figure P1.34



35. The graph in Figure P1.35 shows a relationship between two quantities: position x, measured in meters, and time t, measured in seconds. Describe the relationship using (*a*) a verbal expression and (*b*) a mathematical expression. ●●●

Figure P1.35



36. You have four equilateral triangles, two red and two blue. You want to arrange them so that each blue triangle shares a side with the other blue triangle and with both red triangles, and each red triangle shares a side with the other red triangle and with both blue triangles. How can this be done? ●●●

1.6 Physical quantities and units

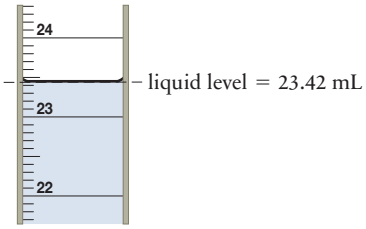
37. The world record for the 100 m dash was set at 9.58 seconds a few years ago. What is this speed in kilometers per hour? In miles per hour? ●
38. A weather balloon can rise to an altitude of 80,000 ft above sea level. What is this distance in miles? In meters? ●
39. A metal block of uniform mass density is cut into two pieces. How does the density of each piece compare with the density of the original block (*a*) if each piece has half the volume of the original block and (*b*) if one piece has one-third of the original volume? ●
40. Take the speed of sound in air at standard temperature and pressure to be 343 *m/s*. Convert this speed to kilometers per hour, inches per nanoseconds, and miles per second. ●●
41. Suppose you measure the mass and volume of two solid stones. Stone 1 has a mass of 2.90×10^{-2} kg and a volume of 10.0 cm^3. Stone 2 has a mass of 2.50×10^{-2} kg and a volume of 7.50 cm^3. Is it likely that the stones are made from the same material? ●●
42. The Dubai World Cup horserace is 2000 meters long. Express this distance in miles. ●●
43. You measure a quantity T that can be represented as $= 2\pi\sqrt{\frac{l}{g}}$. If T and l have the units of seconds and meters, respectively, then what should the units of g be for the expression to be correct? ●●
44. Both *mass* and *volume* can be used to describe the amount of matter present in an object. What physical quantity relates these two concepts? ●●
45. The average radius of Earth is 6371 km. Give order-of-magnitude estimates of what the mass of Earth would be if the planet has the mass density of (*a*) air (≈ 1.2 kg/m^3), (*b*) 5515 kg/m^3, and (*c*) an atomic nucleus ($\approx 10^{18}$ kg/m^3). ●●
46. Suppose $x = ay^{3/2}$, where $a = 7.81\ \mu g/Tm$. Determine the value of y when $x = 61.7\ (Eg \cdot fm^2)/(ms^3)$. Express the result in scientific notation and simplify the units. (Hint: Refer to Table 1.3, and note that SI prefixes are never used to multiply powers of units. For example, the abbreviation cm^2 means $(10^{-2}\ m)^2$, not $10^{-2}\ m^2$, and ns^{-1} is 1/ns or $10^9\ s^{-1}$, not $10^{-9}\ s^{-1}$. Also note that m can stand for meter or for milli, depending on the context.) ●●●

1.7 Significant digits

47. (*a*) What is the speed of light in vacuum expressed to three significant digits? (*b*) What is the square of the speed of light in vacuum expressed to three significant digits? (*c*) Is the numerical value in part *b* the square of the numerical value in part *a*? Why or why not? ●
48. Express the distance of the Dubai World Cup horserace in kilometers to the same number of significant digits needed to answer problem 42. ●
49. You have 2.28 kg of rice and wish to divide it evenly among five people. If you calculate how much rice each person receives, how many significant digits does your answer have? ●
50. You are recording your car's petrol consumption. Your odometer measures to the tenth of a kilometer. The gas-station pump displays liters of gasoline dispensed to the thousandth of a liter. Given these levels of precision, is there any difference in the precision of your calculation when you drive 60.0 kilometers or 600.0 kilometers? ●●
51. The odometer on your car says 35,987.1 km. To the correct number of significant digits, what should the odometer read after you drive another 47.00 m? ●●

52. A certain brand of soft drink contains 34 mg of caffeine per 355-mL serving. If one mole of caffeine has a mass of 194.19 g and you drink an average of two servings of this drink a day, how many caffeine molecules do you ingest in one year? ••
53. You work in a hospital and are preparing a saline solution for a patient. This solution must have a concentration of 0.15 mol of NaCl (mass of one mole, 58.44 gram) per liter of solution. You have lots of volume-measuring devices, but the only mass-measuring tool you have is a balance that measures mass to the nearest 0.1 g. What is the minimum amount of solution you should mix? ••
54. You find a container of an unknown clear liquid. You are interested in its contents, but you have no chemical analysis devices at hand. You do have a good electronic scale, so you measure out 25.403 g of the liquid and then pour this liquid into a graduated cylinder, with the level indicated in Figure P1.54. What is the mass density of this liquid? ••

Figure P1.54



55. Suppose that, in a laboratory exercise, you add 0.335 g of liquid to a 145.67-g ceramic bowl. After you heat the liquid for 25.01 s, you measure the mass of the bowl and liquid together to be 145.82 g. What was the average mass of liquid that evaporated each second? •••

1.8 Solving problems

56. A test question gives the mass density of seawater and asks what volume is occupied by 1.0×10^3 kg of seawater. Someone answers 0.9843 m. Without knowing the mass density value given, explain what is wrong with this answer. •
57. You measure a watch's hour and minute hands to be 8.0 mm and 11.3 mm long, respectively. In one day, by how much does the distance traveled by the tip of the minute hand exceed the distance traveled by the tip of the hour hand? ••
58. An Olympic running track consists of two straight sections, each 84.39 m long, and two semicircular ends, each with a radius of 36.80 m as measured from the inside lane (lane 1). There are eight lanes, each 1.22 m wide. If runners did not start staggered (so that each runner covers the same distance), by how much would the distance traveled by a runner in lane 8 exceed the distance traveled by a runner in lane 1 over one lap? ••
59. You have three coins that are identical in appearance. One is counterfeit, and its mass is different from the mass of the two legitimate coins. You also have an equal-arm balance (a device that compares the masses of two objects and "balances" if they are equal). Describe how you can determine which is the counterfeit coin and whether it is heavier or lighter than the legitimate coins by making only two measurements with the balance. ••
60. A certain type of rice has an average grain length and diameter of 6 mm and 2 mm, respectively. One cup of this rice, after being cooked, contains about 785 food calories. (*a*) How many grains are there in 1 cup of rice? (*b*) How many calories are there in one grain? (*c*) How many cups of uncooked rice are needed to provide 2000 food calories to each of four adults? •••

1.9 Developing a feel

61. What is the order of magnitude of the number of moles of matter in the observable universe? •
62. You add about 10^2 raisins to a loaf of bread you are baking. When fully baked, the loaf is hemispherical with a radius of about 8 in. What is the order of magnitude of the average distance between neighboring raisins? •
63. A tree cut down for lumber has a diameter of 0.80 m over its length of 32 m. About how many 6-ft-long, 2-in. by 2-in. boards can be made from this tree? •
64. Estimate the order of magnitude of the number of times the letter d occurs in this text. ••
65. Estimate the combined length of all the hairs on the head of a person who has a a full head of shoulder-length hair. ••
66. An urban legend making the rounds in quantitative reasoning classes is that a student paid for four years at a private university in the United States exclusively with redeemed deposits on soda cans collected from people's trash. Estimate whether this is possible. ••
67. Computer hard drives have increased dramatically in capacity since their introduction in the 1950s. The first commercial hard drive had a capacity of about 3.8 MB stored on 50 circular platters, each 610 mm in diameter. In the early part of this century, hard drives approached 1.0 TB of storage using just five platters, each with a diameter of about 90 mm. By what order-of-magnitude factor does the storage capacity per unit area of a 21st-century drive exceed that of an original drive? ••

Additional Problems

68. At present, the tallest man-made structure is the Burj Khalifa in Dubai. Express the height of this building in nanometers. How many atoms would need to be stacked to get to the same height? •
69. A 200-sheet stack of loose-leaf paper is 2.75 in. thick. What is the thickness in millimeters of one sheet? •
70. A water tank is filled to its 10×10^4-L capacity. (*a*) What is the tank's volume in cubic meters? (*b*) What is the mass of the water in micrograms? (*c*) If your household consumed ten 500-cc bottles of water a day, how long would the water in the tank last? ••
71. You are working on an exhibit of the atom at a science museum. You want the exhibit to have a model of the atom in which the nucleus is big enough for visitors to see, about 500 mm in diameter. The model needs to fit inside a square room with walls about 25 m long. Will it fit if you make the nucleus this size? ••

72. Recent experimental values for the universal gravitational constant, the mass and mean radius of Earth, and the standard gravitational acceleration are

$$G = 6.6738 \times 10^{-11}\ \text{m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$$
$$M_E = 5.9736 \times 10^{24}\ \text{kg}$$
$$R_E = 6.378\,140 \times 10^{6}\ \text{m}$$
$$g = 9.806\,65\ \text{m} \cdot \text{s}^{-2}$$

Assuming a spherical Earth, these quantities are related by the expression $gR_E^2 = GM_E$. (*a*) To how many significant digits is this relation satisfied using the recent data? (*b*) Can you round off all the data to the same number of significant digits before multiplying and still satisfy this relation to that number of significant digits? If so, what numbers of significant digits work? ●●

73. You are in charge of safety at a local park. A tire swing hangs from a tree branch by a 5-m rope, and the bottom of the tire is 1 m from the ground directly beneath it. The land is sloped at a 12° angle toward a pond. From experience, you know that even an energetic user will be unlikely to cause the swing to make a maximum angle exceeding 30° with respect to the vertical. You are certain that parents will want reassurance about the safety of the tire swing. ●●● CR

74. While working for a space agency, you are tasked with ensuring that the oxygen needs of human beings on long voyages in outer space are met. Your boss hands you a data table and suggests that you consider the amount of oxygen a person might need in 1 year as well as the storage requirements for this oxygen at room temperature and atmospheric pressure. ●●● CR

Volume of one breath	4.5 L
Oxygen content of air, by volume	20.95%
Percent oxygen absorbed per breath	25%
Average breathing rate	15 breaths/min
Mass density of air at room temperature and atmospheric pressure	1.0 kg/m^3

75. Neutron stars can have a mass on the order of 10^{30} kg, but they have relatively small radii, on the order of tens of kilometers. (*a*) What order of magnitude is the mass density of a neutron star? (*b*) How many orders of magnitude larger is this than the mass density of Earth? Of water? (*c*) If water had the mass density of a neutron star, what order of magnitude of mass would be contained in the soda in a full 2-L bottle? ●●●



2 Motion in One Dimension

2.1 From reality to model
2.2 Position and displacement
2.3 Representing motion
2.4 Average speed and average velocity

2.5 Scalars and vectors
2.6 Position and displacement vectors
2.7 Velocity as a vector
2.8 Motion at constant velocity
2.9 Instantaneous velocity

CONCEPTS

QUANTITATIVE TOOLS

The branch of physics that deals with the quantitative representation of motion is called *kinematics*. Kinematics does not consider the causes and effects of motion; its goal is simply to provide a mathematical description of motion. In this chapter, we follow a kinematics approach; that is, we represent motion on a graph and quantify that motion without worrying about what causes it. In Chapter 8 we shall begin to look at the causes of motion, but for now all we care about is the motion while it is happening.

By describing the motion represented in a video, we begin to develop the most basic skill in physics: making simplified representations of real-world situations. You'll see how to use tables, graphs, and mathematical functions to represent quantitative data on motion. Finally, you'll learn the distinction between the terms *position* and *displacement* as well as the distinction between *speed* and *velocity*.

2.1 From reality to model

The opening picture of this chapter shows the *Shinkansen*—the Japanese bullet train—speeding by Mount Fuji. Although the image conveys the idea of motion, it does not contain any quantitative representation of that motion. If I were to ask, "How fast is the train moving?" you wouldn't be able to answer if all you had to go on was this photograph. In order to answer my question, you need a more quantitative representation of the motion.

To analyze the motion of an object, we need to keep track of the object's position at different instants. If the position changes as time passes, the object is in motion. If the position is not changing, the object is said to be **at rest.** As a first step in measuring position, let's analyze the video clip reproduced in **Figure 2.1**, a record of me walking (and you thought physics was going to be dull!). As you can see from the clip, I first walked forward a certain distance, stood still for a short while, and then walked backward.

The imaginary straight line on the ground along which I took my steps is our *reference axis*. My position is determined by my location along this axis relative to some arbitrarily chosen reference point called the *origin*. Giving my position at different instants in time determines my motion in much the same way the video does. Choosing the left edge of the frame as the origin, we can say that in frame 1 the center of my body (in reality, the *image* of my body) is 2.5 mm from that point; a few frames later it is 8.5 mm from the edge (frame 5); another few frames and it is 12.0 mm from the edge (frame 8). In other words, we can describe my motion by specifying my position for every frame, which is what **Table 2.1** does.

The same information is shown graphically in **Figure 2.2**, where my position relative to the left edge is plotted along the vertical axis and the horizontal axis represents the frame number. The formal way of describing this

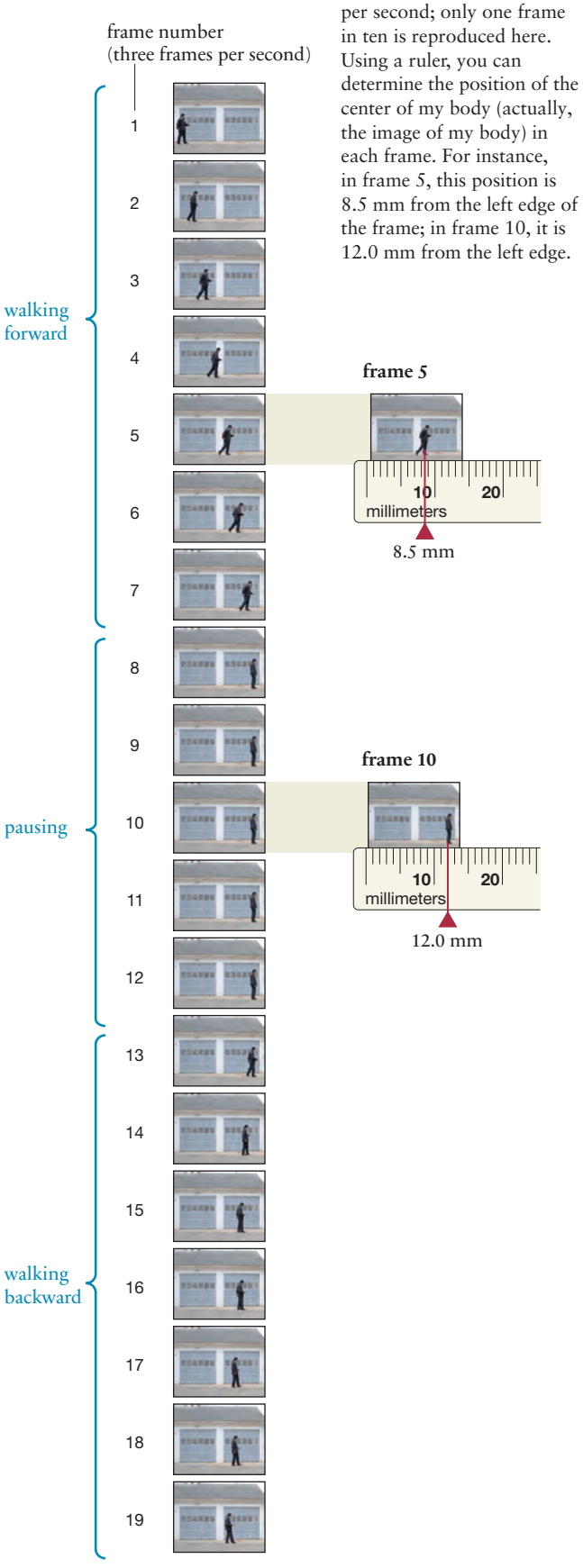


Figure 2.1 Video clip showing me walking to the right, pausing, then walking back to the left. The camera recorded this event at 30 frames per second; only one frame in ten is reproduced here. Using a ruler, you can determine the position of the center of my body (actually, the image of my body) in each frame. For instance, in frame 5, this position is 8.5 mm from the left edge of the frame; in frame 10, it is 12.0 mm from the left edge.

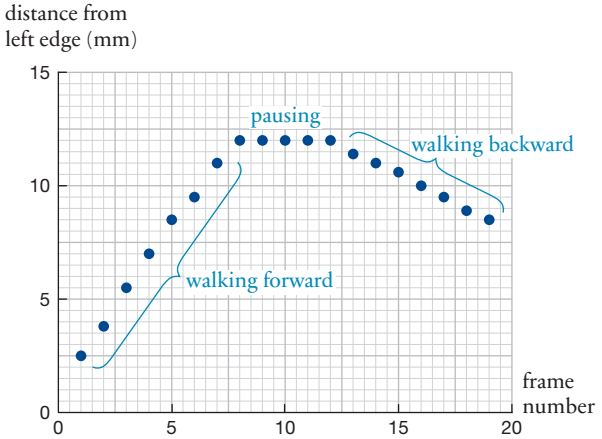
CONCEPTS

Table 2.1 Position versus frame number from the video sequence in Figure 2.1

Frame number	Distance from left edge (mm)	Position x (m)	Time t (s)
1	2.5	+1.0	0
2	3.8	+1.5	0.33
3	5.5	+2.2	0.67
4	7.0	+2.8	1.00
5	8.5	+3.4	1.33
6	9.5	+3.8	1.67
7	11.0	+4.4	2.00
8	12.0	+4.8	2.33
9	12.0	+4.8	2.67
10	12.0	+4.8	3.00
11	12.0	+4.8	3.33
12	12.0	+4.8	3.67
13	11.4	+4.6	4.00
14	11.0	+4.4	4.33
15	10.6	+4.2	4.67
16	10.0	+4.0	5.00
17	9.5	+3.8	5.33
18	8.9	+3.6	5.67
19	8.5	+3.4	6.00

graph is to say that it shows my position (information on the vertical axis) *as a function of* the frame number (information on the horizontal axis). The coordinates of the center of each circle correspond to the values listed in the first two columns in Table 2.1. The size of the circle is representative of the accuracy of the measurements. The upward trend of data points between frames 1 and 8 indicates that my distance from the left edge increases with increasing frame number and represents my motion walking from left to right. Then, representing the time interval when I stood still, the data points stay flat from frame 8 to 12. Finally, from frame 12 onward, the data points show a downward trend representing my motion back to the left.

Figure 2.2 My distance from the left edge of each frame in Figure 2.1 versus frame number. We see that I start, in frame 1, 2.5 mm from the left edge of the frame, walk forward to 12.0 mm from the left edge at frame 8, stay there for five frames, then walk backward to 8.5 mm from the left edge.



2.1 (*a*) Using a metric ruler, measure the position of the center of my body from the left edge in each frame of Figure 2.1. Compare your list of values with the positions given in Table 2.1. (*b*) Plot the positions you measured as a function of frame number—no need for graph paper, just draw your axes and plot your data by eye.

Taking information from the real world and creating a simplified abstract representation of it—making a model of a real-world happening—is one of the most important skills in physics. You'll be making such models hundreds of times throughout this course, some graphical like Figure 2.2, some in the form of tables as Table 2.1, some in the form of mathematical expressions.

2.2 (*a*) In Figure 2.2 the data points for frames 8–12 are all at position 12.0 mm. Explain what this alignment of data points means in terms of what is happening in the real world. (*b*) The data points for frames 7 and 14 are aligned horizontally. What does this alignment imply about the real-world motion? (*c*) What would it mean if two data points were aligned vertically? Could this happen?

2.2 Position and displacement

Figure 2.2 is of limited value as a model of a real-world happening. It tells us only that, in 19 frames of a certain video displayed on a screen, my image moved from a position 2.5 mm from the left edge of the frame to a position 8.5 mm from the edge. It would be more useful to have this information in terms of real-world distances and time intervals. To know what distance I traveled and how long the trip took, I should have put markers on the ground at known intervals and a clock in the camera's field of view. Even though I did not do this, we can still determine position and time by *calibrating* the graph, as you can see in the next checkpoint.

2.3 (*a*) Take a reasonable guess at my height and calculate what real-world distance 1 mm in Figure 2.1 corresponds to. (*b*) Using this result, calculate the distance I walked between frames 1 and 10.

This checkpoint gives a procedure for calibrating the vertical axis in Figure 2.2 without taking any new pictures. Using 1.8 m as my height, I obtained the position values listed in the third column of Table 2.1. For motion in one dimension, position can be represented by the x coordinate along a reference axis (also called the x axis),

CONCEPTS

and this coordinate is denoted by the symbol x. Because the x coordinate can be positive or negative, we always explicitly denote its algebraic sign. In frame 4, for example, my position is given by $x = +2.8$ m. Note that by following this procedure, x values increase to the right.

We can do something similar for time. The camera takes 30 frames per second, and so the time interval between successive frames on the original video is $1/30$ s $= 0.033$ s. If we imagine having started a stopwatch at the first frame, that stopwatch would read 0.033 s at the second frame, 0.067 s at the third, 0.100 s at the fourth, and so on. Because only every tenth frame is shown in Figure 2.1, we must multiply these readings by 10. All these clock readings are listed in the fourth column in Table 2.1. Clock readings are represented by the symbol t; each value of t represents an instant in time. In everyday language *instant* can mean different things. For example, when you say, "I'll be back in an instant," the word *instant* refers to a short time interval. In physics, however, an instant has no duration at all; it represents a single value in time.

Figure 2.3 shows the resulting calibrated data in graphical form, with the position given by x on the vertical axis and time t on the horizontal axis. We refer to such a graph as a *position-versus-time graph*.

To obtain the data shown in Figures 2.2 and 2.3, we arbitrarily chose an origin. An origin is always needed in order to specify a position: The statement "the ball's position is $x = +3.2$ m" makes no sense unless we know from where to start measuring these 3.2 meters. The choice of origin is arbitrary. In analyzing Figure 2.1, for instance, I arbitrarily chose the left edge of the frame as the origin. To show that the choice of origin is arbitrary, suppose we replot our data, but now take my position in frame 1 to correspond to $x = 0$ and the clock reading at the first frame to be $t = 0.33$ s. Doing so produces the graph shown in **Figure 2.4**. Compare this graph with Figure 2.3 and you see that the change in origin and in the instant at which $t = 0$ changes the x and t values of each data point. However, as the next checkpoint shows, changing the origin and the instant at which $t = 0$ does not affect the relative positions of the data points.

Figure 2.3 Position-versus-time graph obtained by calibrating the data points in Figure 2.2. The vertical axis now represents my position, measured in meters, relative to a point on the left edge of the frame. The x and t values are listed also in Table 2.1.

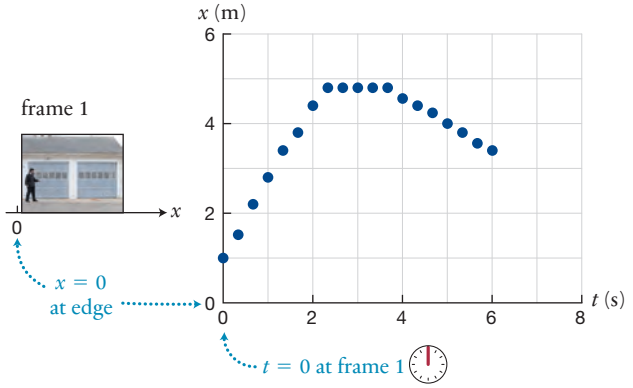
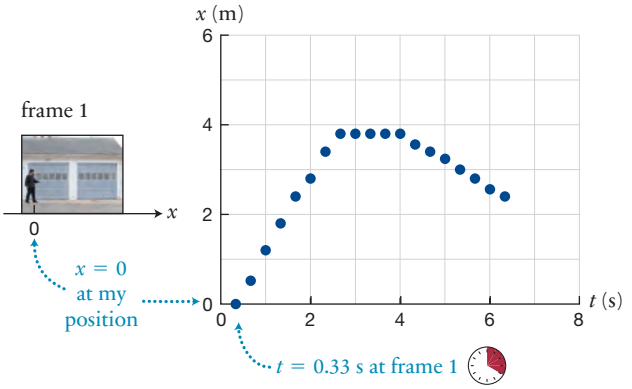


Figure 2.4 Position-versus-time graph for the data plotted in Figure 2.3, but with a different choice of origin.

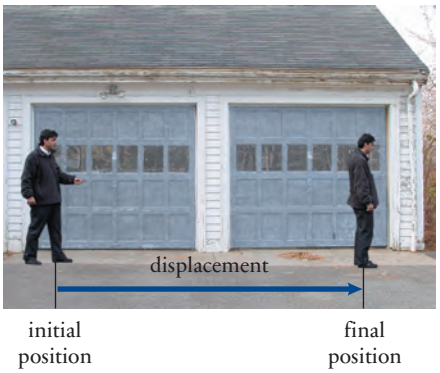


2.4 (*a*) What are the final position values in Figures 2.3 and 2.4? (*b*) What values are obtained by subtracting the initial position value from the final position value in each figure?

We can represent the motion of an object that moves from one position to another by an arrow that points from the initial position to the final position (**Figure 2.5**). This arrow represents a physical quantity called the **displacement** of the object. The displacement does not depend in any way on the choice of reference axis or origin; you don't need any axis to draw the arrow in Figure 2.5. Numerically, we can represent the displacement by subtracting the initial x coordinate from the final x coordinate. The resulting value is called the *x component of the displacement:*

The x component of an object's displacement is the change in its x coordinate. The value of the x component of the displacement is obtained by subtracting the initial x coordinate from the final x coordinate and is independent of the choice of origin.

Figure 2.5 The arrow pointing from my initial position to my final position represents my displacement.



CONCEPTS

The term *x component* serves to remind us that this quantity is measured along some specific x axis. In the second part of this chapter you will see how to specify a displacement like the one in Figure 2.5 by choosing an x axis and determining the x component of the displacement. Later on, in Chapter 10, when we discuss motion that is not along a straight line, we shall need more than one axis and a correspondingly larger number of components to specify displacements.

Note that the x component of the displacement can be either positive or negative: It is positive for displacements in the direction of increasing x and negative for displacements in the direction of decreasing x. To remind ourselves of this fact, we shall always explicitly denote the algebraic sign of the x component of the displacement. The x component of the displacement between my initial and final positions in Figure 2.3, for example, is +2.4 m, indicating that I end up 2.4 m to the right of my starting position.

Returning to Checkpoint 2.4, we see that because every data point in Figure 2.4 is shifted down and to the right by exactly the same amount, the value obtained by subtracting the initial x coordinate from the final x coordinate is not affected by the change in origin. If it puzzles you that the x component of the displacement is independent of the choice of origin, consider the following: Suppose you tell a friend, "See you at the bakery in thirty minutes." Even if your watch says 3:30 at that instant and your friend's watch says 2:30, you will both arrive at the bakery at the same time: 30 minutes after you made the date. Your reference point is the clock reading 3:30 and your friend's reference point is the clock reading 2:30, but both of you add the same value—30 minutes—to your reference point. Just like the x component of the displacement (a *change in* the x coordinate), a time interval (a *change in* the clock reading) does not depend on the choice of reference point.



2.5 Suppose you walk in a straight line from a point P to a point Q, 2 m away from P, and then walk back along the same line to P. (*a*) What is the x component of your displacement for the round trip? (*b*) What distance did you travel during the round trip? (*c*) Is *distance traveled* the same thing as the *x component of the displacement*?

As this checkpoint shows, it is important to distinguish between the x component of the displacement and the *distance traveled*.

The distance traveled is the distance covered by a moving object along the path of its motion.

In contrast to the x component of the displacement, which can be either positive or negative, the distance traveled is always positive. Note also that, for an object that turns around, as in Checkpoint 2.5, the distance traveled is greater than the distance between the endpoints. Only when the motion is always in the same direction is the distance traveled equal to the distance between the endpoints.

2.3 Representing motion

The graphs of Figures 2.3 and 2.4 specify position only at the specific instants represented by the frames in Figure 2.1. What about the positions at intermediate instants? To answer this question, we would need to look at all the frames taken between those shown in Figure 2.1. Doing so would give us ten times as many data points because the camera took ten times the number of frames shown in the figure. If we wanted information more detailed than that, we would need a camera that takes more than 30 frames per second.

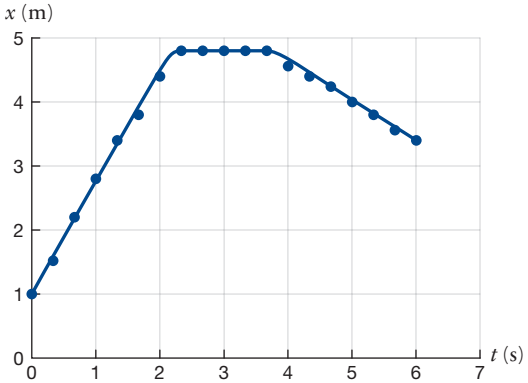
As we repeat our measurements over smaller and smaller time intervals, the spacing between data points becomes increasingly smaller, eventually yielding what looks like a continuous curve instead of discrete points. If we assume the motion proceeds smoothly from frame to frame, we can obtain the same result by *interpolating* the data points in Figure 2.3—in other words, by drawing a smooth curve through the points, as in **Figure 2.6**.



2.6 (*a*) From Figure 2.6, how many seconds did it take me to go from $x = +1.0$ m to $x = +4.0$ m? (*b*) From $x = +2.0$ m to $x = +3.0$ m? (*c*) At what instant did I reach $x = +2.5$ m? (*d*) For how long was I at $x = +2.5$ m?

Figure 2.7 shows how the position-versus-time graph of Figure 2.6 relates to Figure 2.1. Figure 2.6 is a further simplification of our representation of the motion in the video clip. In reality, the motion may have been less smooth than the graph suggests. However, the video clip and Figure 2.6 both tell the same story: I moved forward about 3.8 m in 2.3 s, stood still for 1.3 s, and then moved 1.4 m backward in 2.3 s. Even though it gives less detail than the video clip, the simple curve in Figure 2.6 entirely describes the motion.

Figure 2.6 By interpolating the data points in Figure 2.3, we obtain a smooth curve representing my motion.



CONCEPTS

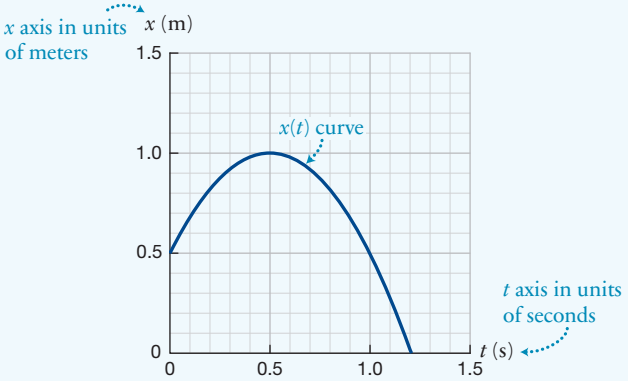
Figure 2.7 If you connect the images in sequential frames by a line, you obtain a curve representing the motion. Turn the page sideways and compare the curve through the images with the curve in Figure 2.6.



EXAMPLE 2.1 Graphical representation

The curve in **Figure 2.8** is a graphical representation of the motion of a certain object. (*a*) What was the x coordinate of the object at $t = 0.50$ s? (*b*) At what instant(s) did the object reach $x = +0.80$ m? (*c*) What distance did the object travel between $t = 0.80$ s and $t = 1.2$ s? (*d*) How long did it take for the object to move from $x = +1.0$ m to $x = 0$?

Figure 2.8 A position-versus-time graph representing the motion of an object. The curve, which shows the position of the object at any instant between $t = 0$ and $t = 1.2$ s, can be represented by a mathematical function $x(t)$ (Examples 2.1 and 2.2).



❶ **GETTING STARTED** This problem requires me to use Figure 2.8 to answer questions about the motion of the object. The problem boils down to reading off values from the graph.

❷ **DEVISE PLAN** To find the object's position x at a given instant t, I draw (or imagine drawing) a vertical line at the value on the t axis corresponding to the given instant in time and read off the value of x at which that line intercepts the curve in the graph. Likewise, to find the instant t at which the object is at a given position x, I draw a horizontal line at the given value of x and read off the value of t at which that horizontal line intercepts the curve.

❸ **EXECUTE PLAN** (*a*) A vertical line drawn at $t = 0.50$ s in Figure 2.8 intercepts the curve at $x = +1.0$ m. ✔

(*b*) A horizontal line at $x = +0.80$ m intercepts the curve at $t = 0.20$ s and at $t = 0.80$ s. ✔

(*c*) At $t = 0.80$ s the object was at $x = +0.80$ m, and at $t = 1.2$ s it was at $x = 0$. The distance traveled was thus 0.80 m. ✔

(*d*) The object was at $x = +1.0$ m at $t = 0.50$ s and at $x = 0$ at $t = 1.2$ s. Thus it took the object 1.2 s − 0.50 s = 0.7 s to move between these points. ✔

✔ The checkmarks show that I carried out the following checks: Vectors and scalars used appropriately; Everything answered; No unknowns left; Units correct; Significant digits justified.

❹ **EVALUATE RESULT** That I obtain two instants in part *b* is OK because the object can return to a certain position. The answers to parts *c* and *d* (0.80 m and 0.7 s) are within the range of the curve (1.0 m and 1.2 s), so my answers are not unreasonable.

CONCEPTS

A curve in a position-versus-time graph like the one shown in Figures 2.6 and 2.8 is called an $x(t)$ *curve* because it can be represented by a mathematical function $x(t)$. The curve and the function specify where the object is at any given instant. Note that $x(t)$ is a single symbol that means "x at instant t"; it is not the product of x and t and should not be confused with the label "x (m)" on the x axis, which means "x in units of meters."

To find the position of the object at a specific instant, we can either read off the position from the graph or substitute that instant into the function. For example, "$x(t = 0.2\text{ s})$" gives the position of the object at the instant $t = 0.2$ s; this position can be obtained by substituting 0.2 s for t in the mathematical expression for $x(t)$.

EXAMPLE 2.2 Mathematical representation

The curve in Figure 2.8 can be mathematically represented by the function $x(t) = a + bt + ct^2$, where $a = 0.50$ m, $b = +2.0$ m/s, and $c = -2.0$ m/s^2. Use this function to answer the four questions in Example 2.1.

❶ **GETTING STARTED** This problem requires me to use a mathematical expression rather than a graph to answer questions about the motion of an object.

❷ **DEVISE PLAN** To find the object's position at a given instant, I substitute that instant for t in the function given in the problem and compute the value of x. To find the instant at which the object is at a given position, I substitute that position for x in the function given in the problem and solve for t.

❸ **EXECUTE PLAN** (*a*) Substituting $t = 0.50$ s into the function, I find for the position at $t = 0.50$ s: $x(t = 0.50\text{ s}) = 0.50\text{ m} + (+2.0\text{ m/s})(0.50\text{ s}) + (-2.0\text{ m/s}^2)(0.50\text{ s})^2 = +1.0$ m. ✔

(*b*) To find the instants at which the object reached $x = +0.80$ m, I must set x in the mathematical function equal to $+0.80$ m: $+0.80\text{ m} = 0.50\text{ m} + (+2.0\text{ m/s})t + (-2.0\text{ m/s}^2)t^2$ or, simplifying the equation by omitting the units, $(2.0)t^2 - (2.0)t + 0.30 = 0$. This is a quadratic equation in t, which has two solutions: $t = [2.0 \pm \sqrt{4.0 - 4 \times 2.0 \times 0.30}]/(4.0)$. Working out this expression yields $t = 0.18$ s and $t = 0.82$ s. ✔

(*c*) At $t = 0.80$ s, the object's position is $x(t = 0.80\text{ s}) = 0.50\text{ m} + (+2.0\text{ m/s})(0.80\text{ s}) + (-2.0\text{ m/s}^2)(0.80\text{ s})^2 = +0.82$ m. At $t = 1.2$ s, the object is at $x = 0.50\text{ m} + (+2.0\text{ m/s})(1.2\text{ s}) + (-2.0\text{ m/s}^2)(1.2\text{ s})^2 = +0.02$ m. The distance it traveled is the difference between these two positions: $x(t = 0.80\text{ s}) - x(t = 1.2\text{ s}) = +0.82\text{ m} - (+0.02\text{ m}) = 0.80$ m. ✔

(*d*) I begin by calculating at what instant the object reached $x = +1.0$ m. Substituting this value into the mathematical expression yields $1.0\text{ m} = 0.50\text{ m} + (+2.0\text{ m/s})t + (-2.0\text{ m/s}^2)t^2$ or, simplifying the equation by omitting the units, $-(2.0)t^2 + (2.0)t + 0.50 = 1.0$. Solving this quadratic equation yields $t = 0.50$ s. Next I must find the instant at which the object reached $x = 0$. Following the same procedure, I find two solutions: $t = -0.20$ s and $t = 1.2$ s. The first solution is before the starting point of the motion represented in Figure 2.8, and so I use the second solution to find that the object took $1.2\text{ s} - 0.50\text{ s} = 0.7$ s to travel this distance. ✔

❹ **EVALUATE RESULT** To check my work, I can compare my answers with the ones I obtained in Example 2.1. Indeed, my answers are the same as (or very close to) the ones I obtained from reading the graph.

2.4 Average speed and average velocity

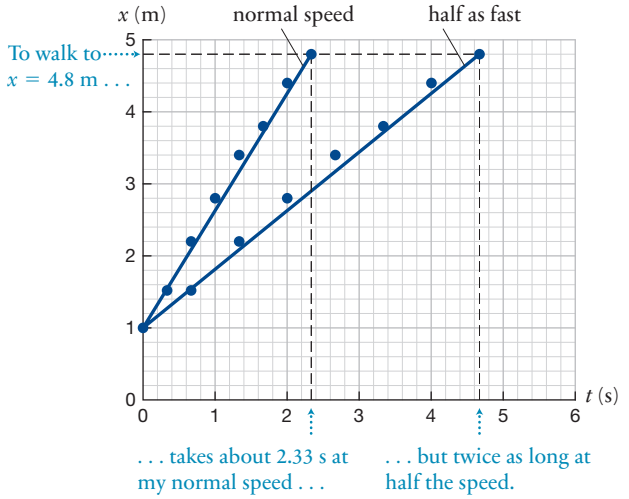
Now that you know how to represent motion in terms of position and time, let us use this information to calculate first the average speed and then the average velocity of the motion.

First, average speed. The more slowly I move, the larger the time interval I require to travel a given distance. Suppose, for instance, I had walked twice as slowly while the camera was rolling. I would have taken twice as much time to get to, say, the end of the forward part of my walk, which is the point $x = +4.8$ m in Figure 2.6. The distance traveled from my starting point would have been the same as before ($4.8\text{ m} - 1.0\text{ m} = 3.8$ m), but it would have taken 4.67 s to reach this point rather than the 2.33 s shown in Figure 2.6. This difference is shown in **Figure 2.9**.

The rate at which an $x(t)$ curve rises with increasing time is called the *slope* of the curve. Notice how the slope becomes less steep when the speed decreases.

2.7 For each speed in Figure 2.9, what are (*a*) the distance traveled between $t = 0$ and $t = 1.50$ s and (*b*) the ratio of the distance traveled to the corresponding time interval? For each speed, what are (*c*) the time interval required to move from $x = 2.00$ m to $x = 3.50$ m and (*d*) the ratio of the distance traveled to the corresponding time interval?

Figure 2.9 Forward motion at two speeds. Walking half as fast means taking twice as long to travel the same distance. To get to the arbitrary point $x = 3.5$ m, for example, takes about 3.0 s at half speed but only about 1.5 s at normal speed.



CONCEPTS

This checkpoint reveals how to extract speed from an $x(t)$ curve: Divide the distance traveled by the time interval required to travel this distance. The ratios of parts *b* and *d* correspond to the slopes of the lines in Figure 2.9 and bear out what I said before: The steeper the slope (the higher the number you got for your ratios), the higher the speed.

Suppose someone rode a bicycle 20 km in 2 h. Using the above procedure to calculate the cyclist's speed, we get $(20\text{ km})/(2\text{ h}) = 10\text{ km/h}$. This doesn't necessarily mean the speed was constant—the person could have, for instance, gone much faster than 10 km/h and taken a break somewhere in the middle of the trip. From the data provided, we can say only that the person's speed was 10 km/h *on average*. Hence:

An object's average speed is the distance traveled divided by the time interval required to travel that distance.

Note that this definition involves the *distance traveled* and not the *x component of the displacement*. To see why, let us again use the data from my walk. According to Table 2.1, the *x* component of my displacement was (final position) − (initial position) = (+3.4 m) − (+1.0 m) = +2.4 m over the 6.0-s duration of the trip. The distance I traveled, on the other hand, is the sum of my forward and backward distances: 4.8 m − 1.0 m = 3.8 m for the forward part and 4.8 − 3.4 = 1.4 m for the backward part. Thus the distance I traveled in 6 s was 5.2 m, so my average speed was (5.2 m)/(6.0 s) = 0.87 m/s, *not* (2.4 m)/(6.0 s) = 0.40 m/s.

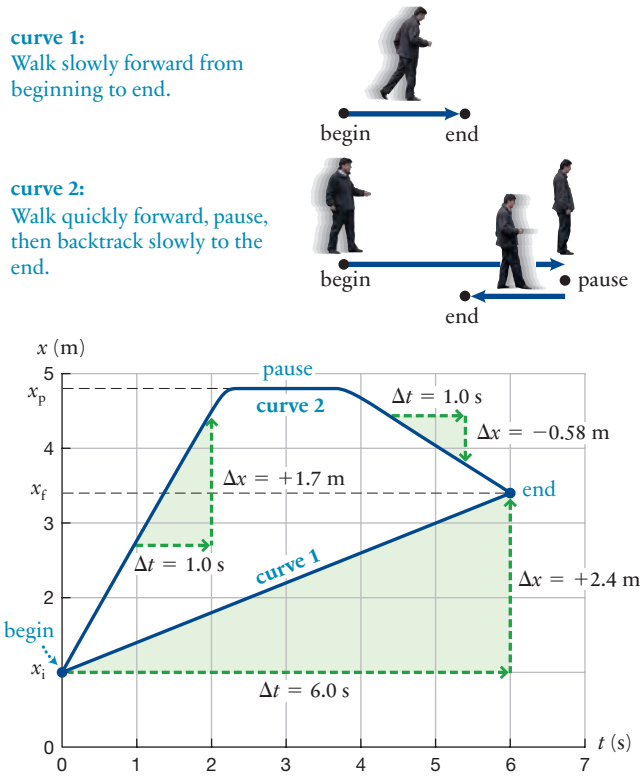
If I had walked at a constant speed of 0.40 m/s for 6 s, my $x(t)$ curve would be curve 1 in **Figure 2.10**. A constant speed of 0.40 m/s implies that I traveled 0.40 m every second. More generally, a constant speed during a given time interval means that when the time interval is divided into any number of equal-sized subintervals, the distance traveled is the same in all subintervals. A constant speed therefore shows up as a straight line in a position-versus-time graph.

Instead of walking at a constant speed from the initial position x_i directly to the final position x_f in Figure 2.10, however, I got to x_f by first walking beyond that point and then walking back, as illustrated by curve 2 in Figure 2.10.

2.8 (*a*) At what instant did I first pass the position $x_f = +3.4$ m (see curve 2 in Figure 2.10)? (*b*) What was my average speed up to that instant?

In Figure 2.10 let us use the symbol x_p to denote the point at which I paused in my walk. Walking forward, I traveled $x_p - x_i = 4.8\text{ m} - 1.0\text{ m} = +3.8$ m in 2.3 s,

Figure 2.10 Two of the infinite number of ways for getting from an initial position x_i to a final position x_f. Curve 1: walking in one direction only; curve 2: walking beyond the endpoint and then backtracking. Both ways result in a displacement of +2.4 m in 6 s.



and so my average speed was $(3.8\text{ m})/(2.3\text{ s}) = 1.7$ m/s. After pausing for 1.3 s, I traveled the distance $x_p - x_f = 4.8\text{ m} - 3.4\text{ m} = 1.4$ m in $6.0\text{ s} - 3.6\text{ s} = 2.4$ s. So the average speed at which I walked backward was $(1.4\text{ m})/(2.4\text{ s}) = 0.58$ m/s. My higher forward speed is shown in Figure 2.10 by the steeper slope of curve 2 between x_i and x_p.

In a graph of position versus time, the steeper the slope of the curve, the higher the speed.

Another point to note about curve 2 in Figure 2.10 is that the curve slopes up from x_i to x_p and down from x_p to x_f. The reason is that I moved in the direction of increasing *x* from x_i to x_p and in the direction of decreasing *x* from x_p to x_f. Average speed, as we defined it, tells us only how fast something moves without indicating the direction in which the motion takes place. In contrast, the slope of an $x(t)$ curve tells us both how fast and in what direction an object is moving. Motion in the positive *x* direction gives a rising $x(t)$ curve and positive slope; motion in the negative *x* direction gives a falling $x(t)$ curve and a negative slope.

The quantity that gives us both the speed and the direction of travel is **velocity.** Like displacement, velocity can be represented by an arrow (more on these arrows and their representation in the second part of

CONCEPTS

this chapter). For now, we define the *x component of the average velocity* as follows:

The *x* component of an object's average velocity is the *x* component of its displacement divided by the length of the time interval taken to travel that displacement.

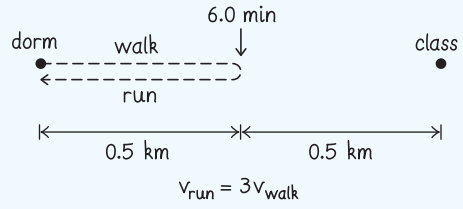
As before, the term *x component* serves to remind us that this quantity depends on the choice of axis and that it can be both positive and negative (in contrast to *average speed*, which is always positive). For motion in the positive *x* direction, the *x* component of the average velocity is positive; for motion in the negative *x* direction, it is negative. Just as we are doing for the *x* component of the displacement, we indicate the direction of motion along the *x* axis by using an algebraic sign. For example, in going from x_i to x_p along curve 2 in Figure 2.10, the *x* component of my average velocity is +1.7 m/s, meaning the motion is in the positive *x* direction; from x_p to x_f it is −0.58 m/s. (My *speed*, however, was 1.7 m/s between x_i and x_p and 0.58 m/s between x_p and x_f. Speed has no direction and therefore carries no algebraic sign.)

EXAMPLE 2.3 Speed and velocity

With about 20 min to spare, you walk leisurely from your dorm to class, which is 1.0 km away. Halfway there, you realize you have forgotten your notebook and run back to your dorm. You walked 6.0 min before turning around, and you travel three times as far per unit of time running as you do when walking. What are (*a*) the *x* component of your average velocity and (*b*) your average speed over the entire trip?

❶ **GETTING STARTED** To begin this problem, it helps to make a sketch to visualize the situation (**Figure 2.11**).

Figure 2.11



❷ **DEVISE PLAN** To find the *x* component of your average velocity, I need to divide the *x* component of your displacement by the time interval over which this displacement took place. To find your average speed, I must divide the distance traveled by the time interval.

❸ **EXECUTE PLAN** (*a*) When you are back at the dorm, the *x* component of your displacement is zero because your initial and final positions are the same. Therefore the *x* component of your average velocity, (*x* component of the displacement)/(time interval), is zero too. ✔

(*b*) Because you traveled halfway to class and back, and the distance between your dorm and the classroom is 1 km, the distance you traveled by the time you returned to your dorm is 1.0 km. You walked for 6.0 min before turning around. Because you run three times as fast as you walk, it takes you only 2.0 min to return to your dorm, and so the total round-trip time is 8.0 min, or 480 s. Your average speed therefore is (distance traveled)/(time interval)=(1000 m)/(480 s)= 2.1 m/s. ✔

❹ **EVALUATE RESULT** If you start and end at the same place, your displacement is zero and so your average velocity is zero, regardless of how fast you move. The answer I obtained in part *b* means that you would cover about (2.1 m/s)(3600 s) = 7600 m = 7.6 km in one hour. This speed—(7.6 km/h)/(1.6 km/mi) = 4.7 mi/h—is in the right ballpark for a person doing a combination of walking and running.



2.9 In Example 2.3, what was your average speed during the time interval that you (*a*) walked away from and (*b*) ran back toward your dorm? (*c*) Is the average of your answers to parts *a* and *b* the same as the average speed calculated in Example 2.3? Why or why not?

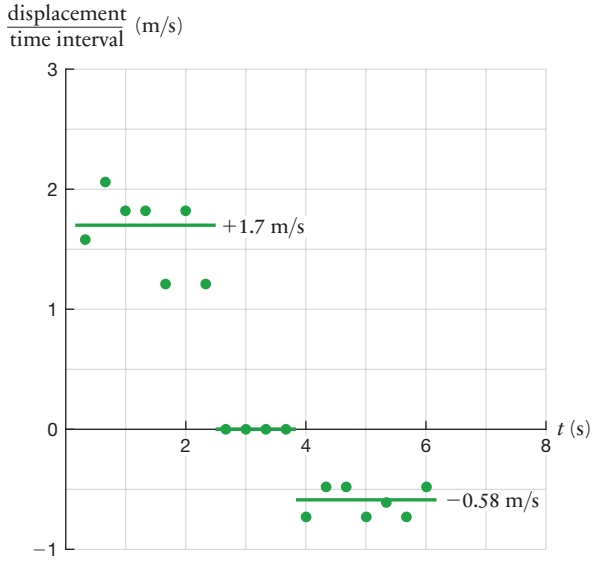
Let us now return to the motion illustrated in the video clip. In Figure 2.10 we determined the *x* component of the average velocity for my motion in each direction: +1.7 m/s walking forward, zero standing still, and −0.58 m/s walking backward. We can obtain more detailed information by calculating the *x* component of the average velocity from one frame to the next. We can do this by calculating the *x* components of the displacements between successive frames from the values of *x* in Table 2.1. The change in the value of *x* from one frame to the next gives the *x* component of the displacements between frames. If we then divide these displacement values by the time interval between frames (0.33 s), we get values for the *x* component of my average velocity that are averaged over much smaller time intervals than before. The result of this computation is shown in **Figure 2.12**.

The first thing that stands out in this graph is that the *x* component of the average velocity obtained from one frame to another varies quite a bit around the *x* component of the average velocity we calculated before. This variation is due to variations in my speed as I walked and to limits in the accuracy of determining the relatively small displacement values from one frame to the next.

Another striking feature in Figure 2.12 is that the *x* component of the average velocity is *negative* for $t \geq 4$ s. This is because the calculation for the *x* component of the displacement (final *x* coordinate minus initial *x* coordinate) automatically takes into account the direction in which the motion takes place. Between $t = 5.0$ s and 5.3 s, for instance, I moved from $x = +4.0$ m to +3.8 m

CONCEPTS

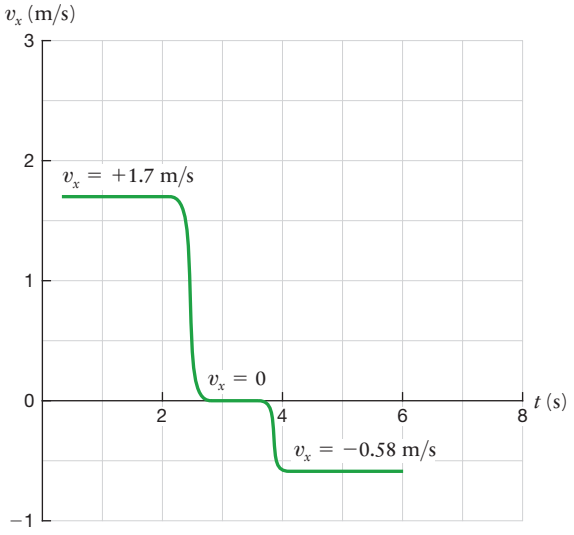
Figure 2.12 Ratio of displacements to time intervals from frame to frame in Figure 2.1.



(see Table 2.1), and so the x component of my displacement was $+3.8\text{ m} - (+4.0\text{ m}) = -0.2\text{ m}$. Whenever x decreases, the x component of the displacement is *negative*. The time intervals, on the other hand, are always positive because time runs forward only. Therefore the ratio of the x component of the displacement to the time interval has the same algebraic sign as the x component of the displacement.

Figure 2.13 shows how the x component of the average velocity, denoted by v_x, changes as a function of time. The curve in this figure (which we'll refer to as a

Figure 2.13 Velocity-versus-time graph for the data shown in Figure 2.12.



$v_x(t)$ *curve*) contains the same information as the $x(t)$ curve in Figure 2.6. Both are simplified representations of the motion shown in Figure 2.1.

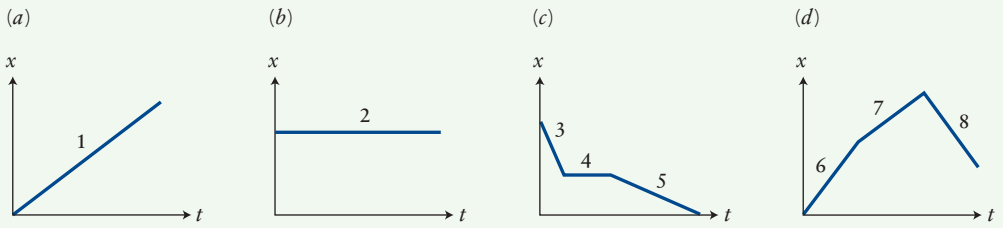
2.10 (*a*) Suppose an object moves from an initial position $x_i = -1.2$ m to a final position $x_f = -2.3$ m without reversing its direction of travel. Is the x component of its average velocity positive or negative? (*b*) Repeat for an object moving from $x_i = +1.2$ m to $x_f = -1.2$ m. (*c*) Do your answers agree with the above statement about the algebraic sign of the x component of the average velocity being the same as the sign of the x component of the displacement?

CONCEPTS

Self-quiz

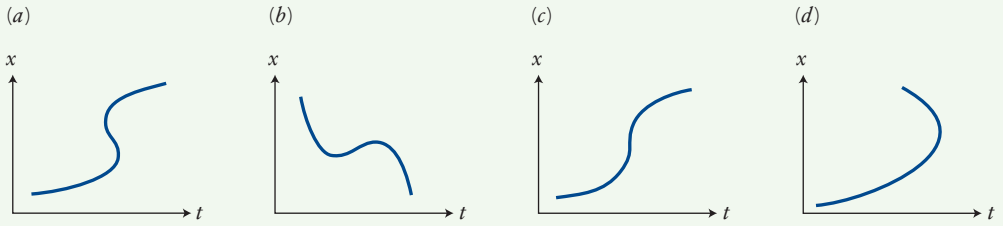
1. Describe the motion represented by each graph in **Figure 2.14**.

Figure 2.14



2. One or more of the graphs in **Figure 2.15** represent an impossible motion. Identify which ones and explain why the motion is not possible.

Figure 2.15



3. Two cars start from some point P, drive 10 km to a point Q, and then return to P. Car 1 completes the trip in 10 min; car 2 takes 11 min. For each car, what are (*a*) the x component of the average velocity and (*b*) the average speed?

4. Suppose you travel 10 km from P to Q at 10 km/h and another 10 km from Q to R at 20 km/h. What is your average speed during the entire trip from P to R?

5. (*a*) Which of the numbered lines 1–8 in the graphs in Figure 2.14 represents the greatest average speed? (*b*) For which of the numbered lines is the x component of the average velocity negative?

Answers

1. (*a*) An object moving away from the origin at constant speed; (*b*) an object at rest a certain distance from the origin; (*c*) an object moving toward the origin, then pausing halfway between its starting point and the origin, then continuing (more slowly) toward the origin; (*d*) an object moving away from the origin at constant speed for one-third of the time, then continuing in the same direction at a slower constant speed for another third of the time, and finally moving toward the origin at some intermediate constant speed.
2. (*a*) and (*d*). Put the left edge of a ruler along the x axis and slowly slide the ruler to the right. The point where this edge intersects the curve shows the position of the object at the instant indicated by the horizontal axis. For both (*a*) and (*d*) there are instants at which the ruler edge intersects the curve at more than one place for the same value of t. This tells you that these curves are meaningless because no object can be in two places at the same instant.
3. (*a*) The x component of the displacement of both cars is zero because they end where they began. So the x component of their average velocities is zero. (*b*) The distance the two cars travel, however, is 20 km. So the average speed of car 1 is (20 km)/(10 min) = 2.0 km/min, and the average speed of car 2 is (20 km)/(11 min) = 1.8 km/min.
4. If you answered 15 km/h, think again—you cannot simply average the two speeds! To calculate the average speed, you need to know how long the entire trip from P to R took and the distance traveled. The 10 km from P to Q takes (10 km)/(10 km/h) = 1 h; the 10 km from Q to R takes (10 km)/(20 km/h) = 0.5 h, and so the total duration of the trip is 1.5 h. The average speed for the 20-km trip from P to R is thus (20 km)/(1.5 h) = 13 km/h.
5. (*a*) Line 3 because its slope is steepest. (*b*) Lines 3, 5, and 8 because the $x(t)$ curves slope downward with increasing time.

2.5 Scalars and vectors

The physical quantities we use in this book fall into two categories. The first type, called **scalars,** are completely specified by a number, which can be positive or negative, and a unit of measure. An example is temperature, a quantity that is specified completely by a number and a unit (as in 21 °C or 70 °F). The second type, called **vectors,** require, in order to be completely specified, both a magnitude (the "how much" and "of what," specified by a number and a unit) and a direction in space.

EXERCISE 2.4 Scalar or vector?

(*a*) Is the statement *Yesterday the car was parked 3 m from the telephone pole* sufficient to determine the position of the car? Is position a scalar or a vector? (*b*) Is the statement *New York is 200 miles from Boston* sufficient to determine the distance between the two cities? Is distance a vector or a scalar? (*c*) Is the statement *My friend walked 3 m due north from the telephone pole* sufficient to determine your friend's position at the end of her short walk? Is displacement a vector or a scalar?

SOLUTION (*a*) No. The car could have been anywhere on the perimeter of a circle of radius 3 m centered on the telephone pole. A direction is required in addition to a distance, and so position is a vector. ✔

(*b*) Yes. All that is required to specify distance is a number (200) and a unit of length (miles). Because distance does not depend on a direction in space, it is a scalar. ✔

(*c*) Yes. *3 m north* completely specifies your friend's displacement from the telephone pole because it gives both a distance (3 m) and a direction (north). Because it requires a direction, displacement is a vector. ✔



2.11 Are the following quantities vectors or scalars: (*i*) the price of a movie ticket, (*ii*) the average velocity of a ball launched vertically upward, (*iii*) the position of the corner of a rectangle, (*iv*) the length of a side of that rectangle?

Scalars follow the ordinary arithmetic laws of addition, subtraction, multiplication, and division. For example, 2 m + 3 m = 5 m and 2 × 5 s = 10 s. Vectors, on the other hand, obey special laws of addition, subtraction, and multiplication that we begin to discuss in the following sections and then learn about in more detail in Chapter 10.

The only type of motion we study through Chapter 9 is motion in one dimension, which means it occurs either forward or backward along a straight line. For now, therefore, the difference between scalars and vectors is not a major issue because in one dimension the direction of any vector can be specified by an algebraic sign. The difference between scalars and vectors becomes important when we discuss motion in more than one dimension in Chapter 10. It is therefore important that we be able to distinguish vectors from scalars, and so in this book we use arrows to denote vectors:*

$\vec{a}$ vector

b scalar.

The arrow is a reminder that vectors have a direction in space and that they obey rules that are different from those for scalars. Throughout this book, we graphically represent vectors by thick arrows of appropriate

*Note that the direction of the arrow used in the vector notation (that is, the $\rightarrow$ in $\vec{a}$) does not represent the direction of the vector. In particular, it does not reverse direction when we multiply a vector by −1.

length and direction and label these arrows with the corresponding symbol (see, for example, Figures 2.16 and 2.17).

The **magnitude of a vector** tells us how much there is of that vector. In Exercise 2.4*c*, for instance, the magnitude of your friend's displacement is 3 m. When we draw one vector arrow longer than another in a diagram, we are saying that the vector represented by the longer arrow has a magnitude that is greater than the magnitude of the vector represented by the shorter arrow. A magnitude never has a sign associated with it.

There are two common ways of designating the magnitude of a vector symbolically: by drawing vertical lines on either side of its symbol or by removing the overarrow. Thus the magnitude of vector $\vec{b}$ is denoted by either $|\vec{b}|$ or b. In this book, we usually use b for simplicity.

To specify vectors mathematically, we introduce a **unit vector**—a vector whose sole purpose is to define a direction in space. Unit vectors have no units, and the number associated with them (their *magnitude*) is one. They are denoted by a caret or "hat" (ˆ) to distinguish them from ordinary vectors, whose magnitude can have any value; for example, the unit vector pointing in the direction of increasing x along the x axis is denoted by $\hat{\imath}$ (pronounced *i-hat*). In **Figure 2.16*a*** this unit vector is represented by a black arrow that points in the direction of increasing x along the axis. Because unit vectors have a magnitude of 1 and no units, we have

$$|\hat{\imath}| \equiv 1. \tag{2.1}$$

Any vector along the x axis can be written in **unit vector notation** as the product of a scalar and this unit vector:

$$\vec{b} \equiv b_x \hat{\imath} \quad \text{(one dimension)}, \tag{2.2}$$

where b_x is a scalar called the *x component of the vector* $\vec{b}$. The subscript x indicates that the quantity refers to the x axis and reminds you that the sign of b_x depends on our choice of axis direction. If the vector $\vec{b}$ points in the positive x direction (that is, in the direction of $\hat{\imath}$, the direction of increasing x), the scalar b_x is positive. If the vector $\vec{b}$ points in the negative x direction (that is, opposite the direction of $\hat{\imath}$, the direction of decreasing x), the scalar b_x is negative.

If we take the absolute value of Eq. 2.2 and substitute Eq. 2.1, we see that, for vectors in one dimension, the magnitude of the vector $\vec{b}$ is equal to the absolute value of b_x:

$$b \equiv |\vec{b}| = |b_x \hat{\imath}| = |b_x||\hat{\imath}| = |b_x| \quad \text{(one dimension)}. \tag{2.3}$$

For practice, let us determine the x components of the vectors in Figure 2.16 and also write the vectors in unit vector notation. In Figure 2.16*b*, the vector $\vec{b}$ points in the positive x direction, which means that the scalar b_x (which is the x component of $\vec{b}$) is a positive number. The length of the arrow in Figure 2.16*b* tells us that the magnitude of $\vec{b}$ is 2.0 m (no sign; never a sign for a *magnitude*). From Eq. 2.3, therefore, the absolute value of b_x is 2.0 m. Consequently, the x component of $\vec{b}$ is $b_x = +2.0$ m. In unit vector notation, we write $(+2.0\text{ m})\hat{\imath}$ for our vector $\vec{b}$.

In Figure 2.16*c*, the vector $\vec{c}$ points in the negative x direction, and so the scalar c_x (the x component of c) is a negative number. The magnitude of $\vec{c}$ along the x axis is 1.5 m, which means the absolute value of c_x is 1.5 m. The x component of $\vec{c}$ is thus $c_x = -1.5$ m. In unit vector notation, we write $(-1.5\text{ m})\hat{\imath}$ for our vector $\vec{c}$.

Figure 2.17 summarizes how a vector $\vec{b}$ is represented graphically and mathematically. To remind ourselves that x components are signed numbers, we always show the algebraic sign of a component, even if it is positive

Figure 2.16 The product of a unit vector $\hat{\imath}$ and a scalar b_x yields a vector $\vec{b} = b_x \hat{\imath}$.

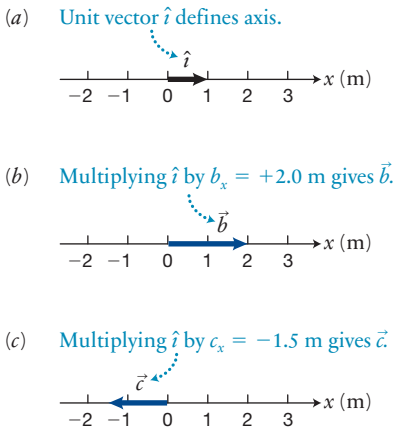
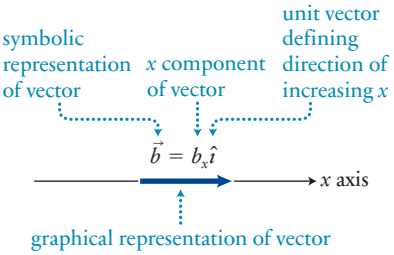
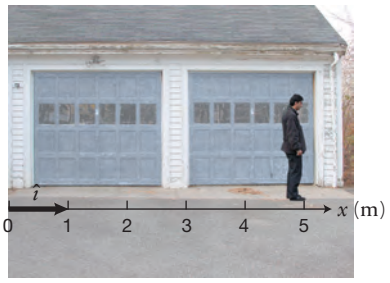


Figure 2.17 Graphical and mathematical representation of a vector.



- magnitude of vector: $b = |b_x|$
- If $b_x > 0$: $\vec{b}$ is in same direction as $\hat{\imath}$
- If $b_x < 0$: $\vec{b}$ is in opposite direction to $\hat{\imath}$

Figure 2.18 The unit vector $\hat{\imath}$ completely specifies a reference axis. The unit vector's direction determines the direction of increasing x along the axis. The tail of the vector locates the origin, and its length defines the unit.



(for instance, $b_x = +2.0$ m). Note also that the x component of a vector always has the same units as the quantity represented by that vector.

In Chapter 10 you'll learn how to mathematically describe vectors that do not lie along the x axis. For motion in one dimension, however, we can always align the axis along the line of motion, and so all vectors can be written in the form given by Eq. 2.2. Because x components are sufficient to completely specify vectors in one dimension, we shall for now work with mostly components rather than vectors.

2.12 Consider the axis and unit vector shown in **Figure 2.18**. Let the vector $\vec{r}$ have its tail at the origin and its tip at my feet. (*a*) What is the x component of that vector? (*b*) Write the vector $\vec{r}$ in terms of its x component and the unit vector. (*c*) What does the vector $\vec{r}$ represent? (*d*) Does $r_x = +2.5$ m unambiguously determine a vector in Figure 2.18? If so, draw the vector, and state what it might represent.

2.6 Position and displacement vectors

Figure 2.19 Superposition of frames 3 and 10 of Figure 2.1. The arrow is my displacement.



The most fundamental vector is displacement, which, as noted in Section 2.2, can be represented graphically by an arrow that points from an object's initial position to its final position. **Figure 2.19**, for example, shows the displacement from my position in frame 3 in Figure 2.1 to my position in frame 10. This vector has a magnitude that is equal to the distance between my initial and final positions (2.6 m), and its direction is toward increasing values of x along the reference axis.

Mathematically we denote displacement by $\Delta\vec{r}$, where the Greek letter Δ (delta) means "the change in" and $\vec{r}$ is a vector specifying position (more on this vector in a bit). The change in a quantity is always taken to be the final value minus the initial value. If your height h increases from 1.75 m to 1.78 m, for instance, the change in your height is $\Delta h = 1.78\text{ m} - 1.75\text{ m} = +0.03$ m. If the number n of cookies in a jar is down to 10 from an initial 12, the change is $\Delta n = 10 - 12 = -2$ cookies.

As we saw in Section 2.2, the x component of the displacement is equal to the change in the x coordinate: $\Delta r_x \equiv \Delta x$. Because of this identity, which is specific to position and displacement, we henceforth denote the x component of the displacement by Δx rather than by Δr_x. So, if the x coordinate of an object changes from some initial value x_i at instant t_i to some final value x_f at some later instant t_f, the x component of the object's displacement is given by

$$\Delta x \equiv x_f - x_i. \tag{2.4}$$

As we saw in Section 2.2, the *x component of the displacement* and the *distance traveled* are not necessarily the same. In particular, if an object turns around during its motion, we need to add the distances between successive turning points (including the initial and final positions) in order to calculate the distance traveled. The **distance** d between two points is equal to the absolute value of the difference in the x coordinates of those two points:

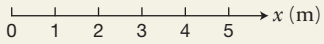
$$d \equiv |x_1 - x_2| \quad \text{(one dimension)}. \tag{2.5}$$

Because distance is obtained by taking the absolute value of a quantity, it is always positive (or zero). Likewise, the distance traveled, which is the sum of the distances between successive points, is always positive (or zero).

EXAMPLE 2.5 Displacement as a vector

On the reference axis shown in Figure 2.20, draw vectors representing my displacements between the following frames in Figure 2.1: (*a*) 4 and 8, (*b*) 11 and 19, (*c*) 6 and 17. Use the values listed in Table 2.1 to describe each vector in terms of its magnitude and direction.

Figure 2.20 Example 2.5.

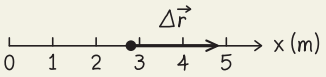


❶ **GETTING STARTED** The pairs of frames given in each part of this example define the initial and final positions. For each pair, I need to determine the displacement vector.

❷ **DEVISE PLAN** I know that a displacement vector points from the initial position to the final position, and so I must draw the tail of the vector at the initial x coordinate and its head at the final x coordinate. Then I can use Eq. 2.3 to determine the magnitude of the displacement vector.

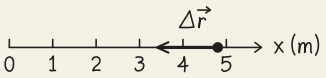
❸ **EXECUTE PLAN** (*a*) The displacement points from my initial position at $x_i = +2.8$ m to my final position at $x_f = +4.8$ m (see Table 2.1). The x component of the displacement is $\Delta x = x_f - x_i = +4.8\text{ m} - (+2.8\text{ m}) = +2.0$ m. The magnitude of this vector is $|\Delta\vec{r}| = |\Delta x| = 2.0$ m; its direction is to the right because the x component is positive (Figure 2.21). ✔

Figure 2.21



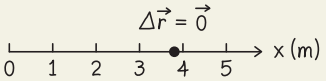
(*b*) The displacement points from $x_i = +4.8$ m to $x_f = +3.4$ m. The x component of the displacement is $\Delta x = x_f - x_i = +3.4\text{ m} - (+4.8\text{ m}) = -1.4$ m. The magnitude of this vector is $|\Delta\vec{r}| = |\Delta x| = |-1.4\text{ m}| = 1.4$ m; its direction is to the left because the x component is negative (Figure 2.22). ✔

Figure 2.22



(*c*) The displacement points from $x_i = +3.8$ m to $x_f = +3.8$ m, and thus it has zero magnitude (and no direction, of course; see Figure 2.23). ✔

Figure 2.23



❹ **EVALUATE RESULT** The arrow in part *a* points right, as I would expect because I moved forward (that is, to the right). In part *b* the arrow points left because I moved backward. Even though I moved between frames 6 and 17, I end at my initial position and so the displacement is zero.

When a vector has zero magnitude as in Example 2.5*c*, we say that it is equal to the *zero vector*, denoted by a zero with an arrow: $\Delta\vec{r} = \vec{0}$. In diagrams, we denote the zero vector by a dot (see Figure 2.23).



2.13 (*a*) What is the x component of the displacement shown in Figure 2.19? (*b*) Write the displacement in terms of its x component and the unit vector.

Rearranging Eq. 2.4, I get an expression for the x coordinate of the final position in terms of the x coordinate of the initial position and the x component of the displacement:

$$x_f = x_i + \Delta x. \tag{2.6}$$

This expression tells you that you can determine the x coordinate of an object's final position by adding the x coordinate of the object's initial position and the x component of its displacement.

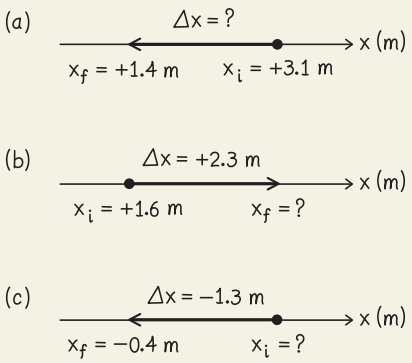
EXAMPLE 2.6 Working with displacements

(*a*) An object moves from an initial position at $x_i = +3.1$ m to a final position at $x_f = +1.4$ m. What is the x component of the object's displacement? (*b*) The x component of an object's displacement is $+2.3$ m. If the object's initial position is at $x_i = +1.6$ m, what is the x coordinate of its final position? (*c*) After undergoing a displacement of -1.3 m, an object is at $x_f = -0.4$ m. What is the x coordinate of the object's initial position?

❶ **GETTING STARTED** To visualize what is happening, I begin by making a sketch showing what is given for each situation and which quantities I need to determine (Figure 2.24).

❷ **DEVISE PLAN** The relationship among the initial x coordinate, the final x coordinate, and the x component of the

Figure 2.24



(Continued)

displacement is given by Eqs. 2.4 and 2.6. For each part of the problem, I know two of the three terms in each of these equations, and so I can use them to find the unknown quantity.

3 EXECUTE PLAN (*a*) The x component of the displacement is given by Eq. 2.4:

$$\Delta x = x_f - x_i = +1.4\text{ m} - (+3.1\text{ m}) = -1.7\text{ m}. \checkmark$$

(*b*) Using Eq. 2.6, I get

$$x_f = x_i + \Delta x = +1.6\text{ m} + (+2.3\text{ m}) = +3.9\text{ m}. \checkmark$$

(*c*) Rearranging terms in Eq. 2.4, I have

$$x_i = x_f - \Delta x = -0.4\text{ m} - (-1.3\text{ m}) = +0.9\text{ m}. \checkmark$$

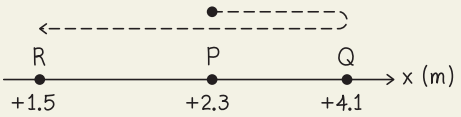
4 EVALUATE RESULT In the sketch I made, the displacement in part *a* points to the left, so the x component of the displacement should indeed be negative. In part *b* the displacement is to the right, so the final x coordinate should be larger than the initial x coordinate. For part *c* I expect the initial x coordinate to be positive because the 1.3-m displacement is larger than the distance from the final x coordinate to the origin (0.4 m).

EXAMPLE 2.7 Displacement, distance, distance traveled

An object moves from point P at $x = +2.3$ m to point Q at $x = +4.1$ m and then to point R at $x = +1.5$ m. (*a*) What is the x component of the object's displacement after traveling from P to R? (*b*) What is the distance between the object's initial and final positions? (*c*) What is the distance traveled by the object?

1 GETTING STARTED A simple sketch helps me visualize the motion and the relative locations of P, Q, and R (Figure 2.25).

Figure 2.25



2 DEVISE PLAN For part *a*, I can use Eq. 2.4 to calculate the x component of the object's displacement given x_i and x_f. In parts *b* and *c*, I need to calculate distances, which are given by Eq. 2.5.

3 EXECUTE PLAN (*a*) The object's initial position is at $x_i = +2.3$ m and its final position is at $x_f = +1.5$ m, so the x component of its displacement is (see Eq. 2.4):

$$\Delta x = x_f - x_i = +1.5\text{ m} - (+2.3\text{ m}) = -0.8\text{ m}. \checkmark$$

(*b*) The distance between the initial and final positions is $d = |x_P - x_R| = |+2.3\text{ m} - (+1.5\text{ m})| = 0.8$ m (see Eq. 2.5). ✔

(*c*) The distance between P and Q is $d = |x_P - x_Q| = |+2.3\text{ m} - (+4.1\text{ m})| = 1.8$ m; the distance between Q and R is $d = |x_Q - x_R| = |+4.1\text{ m} - (+1.5\text{ m})| = 2.6$ m. The distance traveled is thus 1.8 m + 2.6 m = 4.4 m. ✔

4 EVALUATE RESULT The sketch shows that the displacement from P to R is in the negative x direction, and so my answer to part *a* should be negative. The distance between P and R is the absolute value of the displacement, and my answer to part *b* is indeed the absolute value of my answer to part *a*. In part *c*, the total distance traveled (4.4 m) is greater than the distance between P and R. The sketch I made readily confirms that this should be the case.



2.14 (*a*) Consider an x axis on which positive values of x are to the right of the origin. For an object initially located to the left of the origin, is the x component of a displacement to the right positive or negative? (*b*) Does the answer to part *a* depend on whether or not the object crosses the origin?

If displacement is a vector, then the position of an object can also be represented by a vector. To see this, multiply Eq. 2.4 through by $\hat{\imath}$:

$$\Delta x\,\hat{\imath} = (x_f - x_i)\hat{\imath} = x_f\,\hat{\imath} - x_i\,\hat{\imath}. \tag{2.7}$$

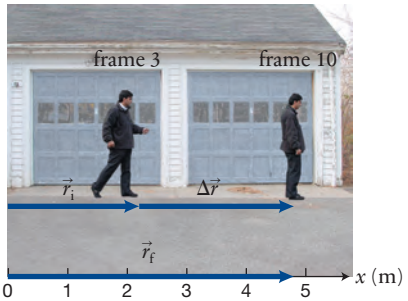
The left side represents the displacement

$$\Delta x\,\hat{\imath} = \Delta r_x\,\hat{\imath} = \Delta\vec{r}. \tag{2.8}$$

The vector $x_i\,\hat{\imath}$ in Eq. 2.7 is represented in Figure 2.26 by an arrow running from the origin to my initial position. This arrow represents the *position vector* (or simply the **position**)—a vector that lets us determine the position of a point in space relative to some chosen origin. Mathematically, position is represented by the symbol $\vec{r}$. For one-dimensional motion we have

$$\vec{r} \equiv x\,\hat{\imath} \quad \text{(one dimension)}, \tag{2.9}$$

Figure 2.26 Superposition of frames 3 and 10 of Figure 2.1. The arrows show my initial and final positions, $\vec{r}_i$ and $\vec{r}_f$, and the relationship of these vectors to the displacement $\Delta\vec{r}$.



QUANTITATIVE TOOLS

and so we see that the x component of position is equal to the x coordinate: $r_x \equiv x$.

My initial position can be thus written as $x_i\,\hat{\imath} = \vec{r}_i$; my final position is $\vec{r}_f = x_f\,\hat{\imath}$, and so Eq. 2.7 becomes

$$\Delta\vec{r} = \vec{r}_f - \vec{r}_i. \tag{2.10}$$

This equation states that the displacement (the change in position) is equal to the final position minus the initial position. Figure 2.26 graphically shows the relationship between displacement $\Delta\vec{r}$ and the initial and final positions.

EXERCISE 2.8 Displacement and position vectors

(*a*) Write each displacement in Example 2.5 in terms of the unit vector $\hat{\imath}$ shown in Figure 2.18. (*b*) What is the magnitude of each displacement vector in part *a*? (*c*) Write my final position in Figure 2.26 in terms of the unit vector.

SOLUTION

(*a*) $\Delta\vec{r} = (+2.0\text{ m})\hat{\imath}$

$\Delta\vec{r} = (-1.4\text{ m})\hat{\imath}$

$\Delta\vec{r} = (0)\hat{\imath} = \vec{0}$. ✔

(*b*) $|(+2.0\text{ m})| = 2.0\text{ m}$

$|(-1.4\text{ m})| = 1.4\text{ m}$

$|0| = 0$. ✔

(*c*) $\vec{r}_f = (+4.8\text{ m})\hat{\imath}$. ✔

Rearranging Eq. 2.10, we get

$$\vec{r}_f = \vec{r}_i + \Delta\vec{r}, \tag{2.11}$$

which tells us that the final position is equal to the sum of the initial position and the displacement. Together with Figure 2.26, Eq. 2.11 gives a graphical prescription for adding two vectors:

> **To add two vectors, place the tail of the second vector in the sum at the tip of the first vector; the vector representing the sum runs from the tail of the first vector to the tip of the second.**

Figure 2.27*a* shows this procedure for two arbitrary vectors $\vec{a}$ and $\vec{b}$.

Equation 2.10 also gives us a prescription for subtracting vectors:

$$\Delta\vec{r} = \vec{r}_f - \vec{r}_i = \vec{r}_f + (-1)\vec{r}_i. \tag{2.12}$$

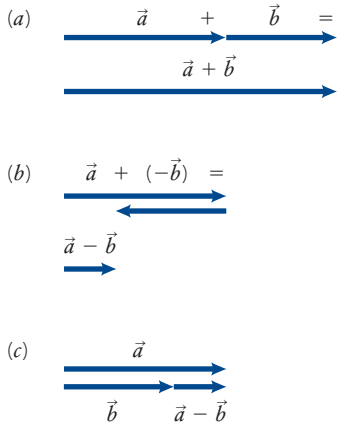
This equation states that we can find the displacement $\Delta\vec{r}$ by subtracting the initial position $\vec{r}_i$ from the final position $\vec{r}_f$ or, alternatively, by adding the vector $(-1)\vec{r}_i$ to $\vec{r}_f$. The vector $(-1)\vec{r}_i$ has the same magnitude as $\vec{r}_i$ but points in the opposite direction. So the graphical prescription for subtracting two vectors is:

> **To subtract a vector from another vector, reverse the direction of the vector being subtracted and add the reversed vector to the vector from which you are subtracting.**

Figure 2.27*b* shows this procedure for two arbitrary vectors $\vec{a}$ and $\vec{b}$.

Figure 2.26 suggests another way of finding the difference between two vectors: The vectors $\vec{r}_i$ and $\vec{r}_f$ both have their tails at the origin, and the

Figure 2.27 (*a*) Adding two vectors. (*b*, *c*) Two ways to subtract one vector from another.



vector $\vec{r}_f - \vec{r}_i$ points from the tip of $\vec{r}_i$ to the tip of $\vec{r}_f$. So an alternative graphical prescription for subtracting two vectors is:

To subtract a vector from another vector, place their tails at the same point and then draw a vector from the tip of the vector being subtracted to the tip of the vector from which you are subtracting.

or

The vector $\vec{a} - \vec{b}$ representing the difference between two vectors $\vec{a}$ and $\vec{b}$ points from the tip of $\vec{b}$ (the vector being subtracted) to the tip of $\vec{a}$.

This approach is illustrated in Figure 2.27*c*.

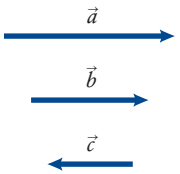
In later chapters we discuss rules for multiplying vectors. For now it suffices to know how to multiply a vector by a scalar. The effect of such a multiplication is the same as that of multiplying the unit vector by a scalar (see Figure 2.16): Multiplying a vector $\vec{a}$ by a scalar c yields a vector whose length equals the length of the original vector times the scalar:

$$c\,\vec{a} = c(a_x\hat{\imath}) = (ca_x)\hat{\imath}. \tag{2.13}$$

If the scalar c is positive, then the product vector points in the same direction as the original vector $\vec{a}$; if c is negative, the product vector points in the direction opposite that of $\vec{a}$.

Although our current treatment of vectors is limited to one dimension, we shall be able to use the same methods when dealing with vectors in more than one dimension in Chapter 10.

Figure 2.28 Checkpoint 2.15.



2.15 Consider the three vectors illustrated in **Figure 2.28**. (*a*) Make a sketch representing the vectors $\vec{a} + \vec{c}$, $\vec{a} - \vec{b}$, and $\vec{c} - \vec{b}$. (*b*) Are the vectors $\vec{a} + \vec{c}$ and $\vec{c} + \vec{a}$ identical? (*c*) Are the vectors $\vec{c} - \vec{b}$ and $\vec{b} - \vec{c}$ identical?

2.7 Velocity as a vector

Like displacement, average velocity is a vector. As we saw in Section 2.4, by dividing the x component of the displacement by the time interval required for that displacement, we obtain the x component of the average velocity during that time interval. Consider, for example, an object for which the x component of the displacement in the time interval between clock readings t_i and t_f is Δx. We can write the time interval as the change in the clock reading, $\Delta t = t_f - t_i$, and so the x component of the average velocity is

$$v_{x,\text{av}} \equiv \frac{\Delta x}{\Delta t} = \frac{x_f - x_i}{t_f - t_i}. \tag{2.14}$$

As before, the subscript x reminds us that this component is measured relative to our reference axis and that it can be either positive or negative, depending on the sign of Δx (Δt is always positive because $t_f > t_i$).

Equation 2.14 gives us the x component of the vector $\vec{v}_{\text{av}} = v_{x,\text{av}}\,\hat{\imath}$. Multiplying the first and middle terms of Eq. 2.14 by $\hat{\imath}$ and substituting $\Delta\vec{r} = \Delta x\,\hat{\imath}$ (Eq. 2.8), we see that average velocity can be written as

$$\vec{v}_{\text{av}} \equiv v_{x,\text{av}}\,\hat{\imath} = \frac{\Delta\vec{r}}{\Delta t}. \tag{2.15}$$

Because Δt is always positive ($t_f > t_i$), Eq. 2.15 shows that for motion in one dimension, the average velocity $\vec{v}_{\text{av}}$ always points in the same direction as the displacement.

QUANTITATIVE TOOLS

EXAMPLE 2.9 Average velocity and average speed

Consider my motion between frames 1 and 8 in Figure 2.1. Use the values listed in Table 2.1 to determine the answers to these questions: (*a*) What is my average speed over this time interval? (*b*) What is the *x* component of my average velocity over this time interval? (*c*) Write the average velocity in terms of the unit vector $\hat{\imath}$.

❶ **GETTING STARTED** Frames 1 through 8 correspond to my forward motion. To determine my average speed and velocity, I need to know both my displacement between these frames and the time interval between them.

❷ **DEVISE PLAN** I can get the average speed from its definition: the distance traveled divided by the time interval. Because I travel in only one direction between frames 1 and 8, the distance traveled is equal to the distance between my initial and final positions. For part *b*, I can use Eq. 2.14 to get the *x* component of the average velocity, and Eq. 2.15 shows me how to write the average velocity in unit vector form.

❸ **EXECUTE PLAN** (*a*) From Table 2.1, $|x_f - x_i| = |+4.8\text{ m} - (+1.0\text{ m})| = 3.8\text{ m}$. I traveled from x_i to x_f in time interval $\Delta t = 2.3\text{ s}$. My average speed was thus $(3.8\text{ m})/(2.3\text{ s}) = 1.7\text{ m/s}$. ✔

(*b*) The *x* component of my displacement is $\Delta x = x_f - x_i = +4.8\text{ m} - (+1.0\text{ m}) = +3.8\text{ m}$. Using Eq. 2.14, I find for the *x* component of average velocity $v_{x,\text{av}} = \Delta x/\Delta t = (+3.8\text{ m})/(2.3\text{ s}) = +1.7\text{ m/s}$. ✔

(*c*) $\vec{v}_{\text{av}} = \Delta\vec{r}/\Delta t = (+3.8\text{ m})\hat{\imath}/(2.3\text{ s}) = (+1.7\text{ m/s})\hat{\imath}$. ✔

❹ **EVALUATE RESULT** The answer to part *a* shows that I cover a distance roughly equal to my height each second, which is reasonable for a person walking. Because I walk in one direction only, the answer to part *b* has the same value as the answer to part *a*. The positive sign indicates that my motion is in the positive *x* direction, as I expect. My answer to part *c* expresses this result in unit vector form.



2.16 (*a*) Consider my motion between frames 6 and 17 in Figure 2.1. Use the values in Table 2.1 to determine the answers to these questions: What is my average speed over this time interval? What is the *x* component of my average velocity? What is the average velocity? (*b*) Repeat for the motion between frames 1 and 17.

2.8 Motion at constant velocity

If an object moves at constant velocity, we can drop the term *average* when referring to velocity. When we do so, the *x* component of the average velocity is simply v_x. According to Eq. 2.14, we thus have

$$\Delta x = v_x\,\Delta t \quad \text{(constant velocity)}. \tag{2.16}$$

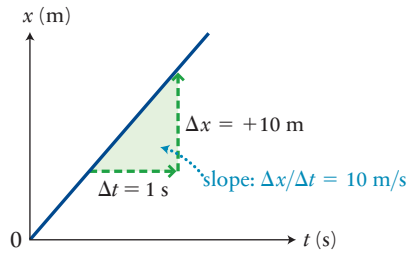
To give just one example, a car going at a constant velocity of +10 m/s in the *x* direction travels a distance of 10 m during each second of its motion. A plot of position versus time is therefore a straight line, as indicated in **Figure 2.29a**. Recall from Checkpoint 2.7 and the text immediately surrounding it that the ratio $\Delta x/\Delta t$ is the slope of the $x(t)$ curve. Looking at Figure 2.29*a*, you can see that this ratio $\Delta x/\Delta t$ is the tangent of the angle that the $x(t)$ curve makes with any horizontal line drawn through the curve. So we see that for motion at constant velocity, the slope of the $x(t)$ curve is numerically equal to the *x* component of the velocity:

$$\frac{\Delta x}{\Delta t} = v_x \quad \text{(constant velocity)}. \tag{2.17}$$

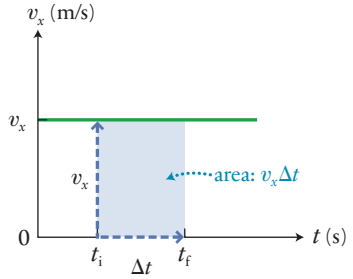
Another useful fact can be obtained by looking at the $v_x(t)$ curve for an object moving at constant velocity. Because the *x* component of the velocity always has the same value when the velocity is constant, this curve is a horizontal line, as shown in Figure 2.29*b*. Notice that the area $v_x\,\Delta t$ (height times width) of the shaded rectangle of width Δt under this line is, from Eq. 2.16, equal to the *x* component of the displacement, Δx, during the time interval.* This is true for any time interval Δt, and so:

Figure 2.29 Graphs of position and velocity as a function of time for an object moving at constant velocity.

(*a*) Position versus time: straight line with nonzero slope



(*b*) Velocity versus time: horizontal line



*The word *area* requires some caution here: The height of the rectangle represents the *x* component of a velocity (in meters per second), and the width represents a time interval (in seconds), and so the product—the "area"—has units of (m/s) × s = m, not m^2 as would a surface area in the traditional sense. This is so because the area of the rectangle, $v_x\,\Delta t = \Delta x$, represents the *x* component of a displacement, which has units of meters.

The x component of the displacement during any time interval is given by the area under the $v_x(t)$ curve between the beginning and the end of the time interval.

Because the x component of the velocity can be positive or negative, the area can also be positive or negative, as it should given that it represents the x component of a displacement. As we shall see in Chapter 3, this correspondence between the x component of the displacement and the area under the $v_x(t)$ curve holds not just for motion at constant velocity but for any type of motion.

For an object moving at constant velocity, Eq. 2.16 gives us a way of determining the x coordinate of the position of the object at any time. Remembering that $\Delta x = x_f - x_i$, we can write Eq. 2.16 in the form

$$x_f - x_i = v_x \Delta t \quad \text{(constant velocity)}, \tag{2.18}$$

or, solving for x_f,

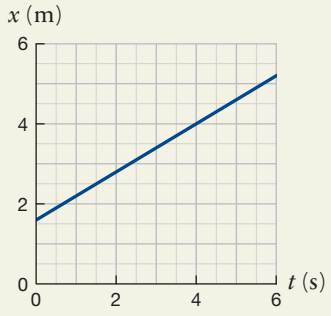
$$x_f = x_i + v_x \Delta t \quad \text{(constant velocity)}. \tag{2.19}$$

This expression allows us to determine, at any time t_f, the x coordinate of the position of an object moving at constant velocity, provided we know the object's initial position coordinate x_i at time t_i.

EXAMPLE 2.10 Motion at constant velocity

Figure 2.30 shows the position-versus-time graph for part of the motion of an object moving at constant velocity. (*a*) What is the x component of the object's velocity? (*b*) Write an expression for $x(t)$, the x coordinate of the position of the object at an arbitrary time t. (*c*) What is the x coordinate of the object's position at $t = 25$ s?

Figure 2.30 Example 2.10.



❶ **GETTING STARTED** The $x(t)$ curve in Figure 2.30 is a straight line, as it should be for a motion at constant velocity. To answer the question, I have to read off information from the graph.

❷ **DEVISE PLAN** The x component of the object's velocity is given by Eq. 2.14. Once I have found this value, I can use Eq. 2.19 to answer part *b*. Although the graph doesn't show the position of the object at $t = 25$ s, I can use my solution to part *b* to find the position by substituting $t = 25$ s.

❸ **EXECUTE PLAN** (*a*) From the graph I see that the object's position changes from $x_i = +1.6$ m to $x_f = +5.2$ m in the time interval between $t_i = 0$ and $t_f = 6.0$ s. Substituting these values into Eq. 2.14, I find for the x component of the velocity:

$$v_x = \frac{x_f - x_i}{t_f - t_i} = \frac{+5.2\text{ m} - (+1.6\text{ m})}{6.0\text{ s} - 0} = \frac{+3.6\text{ m}}{6.0\text{ s}}$$
$$= +0.60\text{ m/s}. \checkmark$$

(*b*) I take the object's position at $t = 0$ as the initial position. Substituting $t_i = 0$, $x_i = +1.6$ m, and $v_x = +0.60$ m/s from part *a* into Eq. 2.19, I get

$$x_f = +1.6\text{ m} + (+0.60\text{ m/s})(t_f - 0)$$
$$= +1.6\text{ m} + (+0.60\text{ m/s})t_f.$$

Because the choice of t_f is arbitrary, I may just as well write, for the x coordinate of the position at any instant of time t:

$$x(t) = +1.6\text{ m} + (+0.60\text{ m/s})t. \checkmark$$

(*c*) Substituting $t = 25$ s into the expression I obtained in part *b* yields

$$x(t = 25\text{ s}) = +1.6\text{ m} + (+0.60\text{ m/s})(25\text{ s})$$
$$= +1.6\text{ m} + 15\text{ m} = +17\text{ m}. \checkmark$$

❹ **EVALUATE RESULT** The x component of the object's velocity is positive, as I would expect from the positive slope of the $x(t)$ curve. I can double-check the value (0.60 m/s) by reading off the increase in x over a 1-s interval in the graph. The result is indeed about 0.6 m. I cannot read off the value for x for $t = 25$ s from the graph because the graph extends to only $t = 6$ s. However, from the graph I see that the object starts near $x = 2$ m and the x coordinate increases by 3 m every 5 s. In 25 s, the object displacement should therefore be $5(3\text{ m}) = 15$ m, and so the object's final position should be near $(2\text{ m}) + (15\text{ m}) = 17$ m, which is what I got.

QUANTITATIVE TOOLS

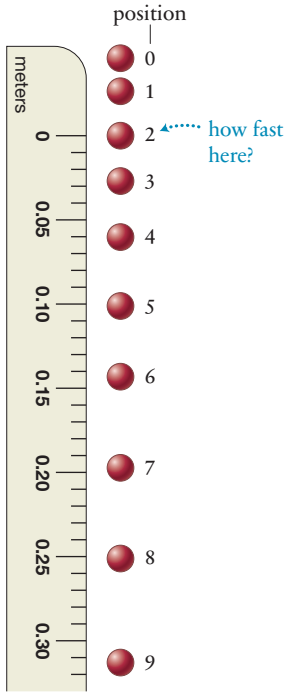
2.17 In Example 2.10, suppose I had taken x_i to be the x coordinate of the object's position not at $t = 0$ but at $t = 2.0$ s. Show that the expression for the x coordinate of the position is the same as the one I obtained in Example 2.10*b*.

2.9 Instantaneous velocity

If you are driving and wonder How fast am I going *now*?, the answer to your question is provided by the speedometer. This gauge gives your *instantaneous speed*, which is generally not equal to your average speed during the entire trip. Likewise, if you drop a ball from a certain height, the ball speeds up as it falls, and so its velocity changes over time. In addition to being able to calculate the ball's average velocity as it falls over a certain distance, you might be interested in knowing its velocity at any given instant—its **instantaneous velocity.**

Figure 2.31 shows successive positions of a falling ball at 0.0300-s time intervals. Such a picture can be obtained using a strobe light to illuminate the ball at fixed intervals and thus record its position at fixed intervals. As the ball falls, the positions at which it is photographed get farther and farther apart. In other words, the magnitude of the ball's displacement per time interval (the magnitude of its average velocity, from Eq. 2.15) increases. Suppose we want to find the ball's velocity at position 2: We want to answer the question How fast is the ball moving at the instant it passes position 2? Using the ruler in the picture, we can measure how long the ball takes to fall a certain distance after position 2, but doing that gives the ball's *average* velocity over that distance. To get a value closer to the ball's instantaneous velocity at position 2, we have to make the time interval over which we measure the distance shorter and shorter.

Figure 2.31 Successive positions of a falling ball recorded at 0.0300-s intervals.



From here on, the word *velocity* without any adjective is used to designate the instantaneous value of the velocity. When we are interested in the average value over a certain time interval, we shall always use the adjective *average*. In equations we drop the subscript av when we mean the instantaneous value—by convention, $\vec{v}_{av}$ = average velocity, $\vec{v}$ = instantaneous velocity. Likewise, *speed* always means *instantaneous speed*.

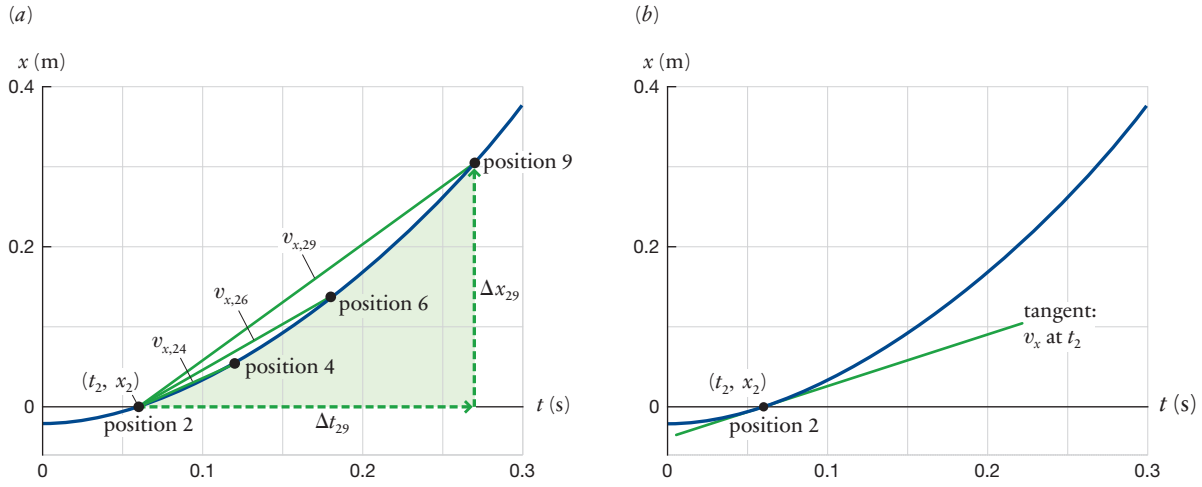
2.18 (*a*) Determine the x component of the ball's average velocity between positions 2 and 9 in Figure 2.31. (*b*) Repeat between positions 2 and 8. (*c*) Repeat for increasingly smaller time intervals. (*d*) As the time interval decreases, does the x component of the average velocity increase, decrease, or stay the same? Explain. (*e*) Plot the x components of the average velocity you obtained in parts *a*–*c* versus the length of the time interval. (*f*) Do you expect the x component of the average velocity to become zero as the time interval approaches zero?

Figure 2.32 shows the position of the ball as a function of time, as obtained from Figure 2.31. The x component of the average velocity measured over an interval Δt is given by Eq. 2.14, and in Checkpoint 2.18 you used that equation to evaluate the x component of the average velocity between positions 2 and 9:

$$v_{x,29} = \frac{x_9 - x_2}{t_9 - t_2}. \qquad (2.20)$$

(I substituted 29 for the subscript av to indicate that we are averaging over the time interval between t_2 and t_9.) This calculation yields the slope of the hypotenuse of the right triangle between positions 2 and 9 in Figure 2.32*a*. As you reduced the size of Δt, you got closer and closer to the value of the velocity at position 2. As the checkpoint shows, the x component of the displacement

Figure 2.32 Procedure for finding the x component of the velocity of the ball in Figure 2.31 from its $x(t)$ curve. (*a*) As we reduce the time interval $\Delta t = t_f - t_2$ over which the x component of the velocity is averaged, Δx shrinks and the straight line between (x_2, t_2) and (x_f, t_f) approaches the tangent to the curve at (x_2, t_2). (*b*) The slope of the tangent at (x_2, t_2) gives the x component of the velocity at that point.



gets smaller as Δt gets smaller, but the ratio of the two (the slope of the hypotenuse) never becomes zero. This is not surprising because the x component of the ball's average velocity had some positive value while passing position 2. Thus that component should also be nonzero at position 2.

Letting the interval Δt approach zero, we obtain the x component of the velocity at instant t:

$$v_x \equiv \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}. \tag{2.21}$$

The right side of this equation is the definition of the *derivative* of x with respect to time:

$$v_x \equiv \frac{dx}{dt}. \tag{2.22}$$

As you can see from Figure 2.32, the hypotenuse of the triangle spanned by the time interval Δt approaches the tangent to the $x(t)$ curve at time t_2, when Δt goes to zero. The x component of the velocity at instant t is thus equal to the slope of the tangent to the $x(t)$ curve at instant t.

The velocity (which is a vector) is then

$$\vec{v} = v_x \hat{\imath} = \left(\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} \right) \hat{\imath}. \tag{2.23}$$

Because the unit vector is constant, we can pull it inside the parentheses and write

$$\left(\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} \right) \hat{\imath} = \lim_{\Delta t \to 0} \frac{\Delta x\, \hat{\imath}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta \vec{r}}{\Delta t}, \tag{2.24}$$

where I have used the relationship $\Delta \vec{r} = \Delta x\, \hat{\imath}$ from Eq. 2.8. The term on the right in Eq. 2.24 is the derivative of the position vector, and so we can write

$$\vec{v} = \frac{d\vec{r}}{dt}. \tag{2.25}$$

The magnitude of this vector is the **speed** $v = |\vec{v}|$.

Table 2.2 Symbols and their meaning

Symbol	Meaning
t	clock reading = instant in time
$\Delta t \equiv t_f - t_i$	change in clock reading = interval of time
x	x coordinate
$d \equiv \lvert x_1 - x_2 \rvert$	straight-line distance between two points (never has sign)
$\hat{\imath}$	unit vector defining direction of x axis (vector magnitude 1)
$\vec{b} \equiv b_x \hat{\imath}$	vector and unit vector notation (see examples below)
b_x	x component of vector $\vec{b}$ (can be positive or negative)
$b \equiv \lvert\vec{b}\rvert = \lvert b_x \rvert$	magnitude of vector $\vec{b}$ (always positive)

Vector	Meaning	Component	Meaning
$\Delta\vec{r} = \Delta r_x \hat{\imath}$	displacement	$\Delta r_x \equiv \Delta x$	x component of displacement = change in x coordinate
$\vec{r} = r_x \hat{\imath}$	position	$r_x \equiv x$	x component of position = x coordinate
$\vec{v}_{av} = v_{x,av} \hat{\imath}$	average velocity	$v_{x,av} \equiv \dfrac{\Delta x}{\Delta t}$	x component of average velocity
$\vec{v} = v_x \hat{\imath}$	(instantaneous) velocity	$v_x \equiv \lim\limits_{\Delta t \to 0} \dfrac{\Delta x}{\Delta t} \equiv \dfrac{dx}{dt}$	x component of (instantaneous) velocity

Table 2.2 summarizes the various symbols we have introduced in this chapter and their meaning.

EXAMPLE 2.11 Time intervals and derivatives

Suppose the x coordinate of the position of an object moving along the x axis varies in time according to the expression $x(t) = ct^3$, where c is a constant. Derive an expression for the x component of the velocity as a function of time.

❶ **GETTING STARTED** There are two ways to obtain the answer to this problem. One is to use differential calculus: The x component of the velocity is the derivative of the function $x(t)$ with respect to time. The other way is to evaluate the x component of the velocity over a finite time interval Δt and then let that interval go to zero. If I use both methods, I can easily verify my answer.

❷ **DEVISE PLAN** I'll use the (harder) second method first. To do so, I have to determine the change in the x coordinates between two instants t_i and $t_f = t_i + \Delta t$. The x component of the average velocity is then the ratio of this change in the x coordinate and Δt. Equation 2.21 then allows me to find the x component of the (instantaneous) velocity at any instant by taking the limit of this result as Δt goes to zero.

❸ **EXECUTE PLAN** The x coordinate of the object's initial position at time t_i is $x_i = ct_i^3$. At a later time $t_f = t_i + \Delta t$, the x coordinate of the position is

$$\begin{aligned} x_f &= ct_f^3 = c(t_i + \Delta t)^3 \\ &= ct_i^3 + 3ct_i^2\Delta t + 3ct_i(\Delta t)^2 + c(\Delta t)^3. \end{aligned}$$

Therefore the x component of the displacement in the time interval Δt is

$$\begin{aligned} \Delta x &= x_f - x_i \\ &= [\,ct_i^3 + 3ct_i^2\Delta t + 3ct_i(\Delta t)^2 + c(\Delta t)^3\,] - ct_i^3 \\ &= 3ct_i^2\Delta t + 3ct_i(\Delta t)^2 + c(\Delta t)^3. \end{aligned}$$

The x component of the average velocity is thus

$$v_{x,av} = \frac{\Delta x}{\Delta t} = 3ct_i^2 + 3ct_i\Delta t + c(\Delta t)^2.$$

If I let Δt approach zero, the second and third terms on the right go to zero, so that I have

$$v_x = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = 3ct^2, \checkmark$$

where I have dropped the subscript i because the result holds for any instant t. This equation gives the velocity at any time t. It tells me that v_x changes as the square of t.

❹ **EVALUATE RESULT** I can verify my result using differential calculus:

$$v_x(t) = \frac{d}{dt}x(t) = \frac{d}{dt}(ct^3) = 3ct^2.$$

EXAMPLE 2.12 The limiting process

Consider again an object moving along the x axis according to the expression $x(t) = ct^3$, with $c = +0.12\ \text{m/s}^3$. (*a*) Determine the x component of the object's average velocity during the interval from $t_i = 1.0$ s to $t_f = 2.0$ s. (*b*) Repeat for the interval from $t_i = 1.0$ s to $t_f = 1.1$ s. (*c*) Show that your result approaches the x component of the velocity at $t = 1.0$ s if you continue to reduce the interval by factors of 10. Ignore the significant digits rules taught in Chapter 1. Instead, use all the digits your calculator shows in every step.

❶ **GETTING STARTED** I have to use a fixed initial instant ($t_i = 1.0$ s) and evaluate the change in the function $x(t)$ over increasingly smaller time intervals Δt, starting with $\Delta t = 1.0$ s.

❷ **DEVISE PLAN** Evaluating the change in $x(t)$ is just a matter of doing it, and so my plan is to roll up my sleeves and dig in!

❸ **EXECUTE PLAN** (*a*) At $t_i = 1.0$ s, the object is at $x_i = (+0.12\ \text{m/s}^3)(1.0\ \text{s})^3 = +0.12$ m. At $t_f = 2.0$ s, it is at $x_f = (+0.12\ \text{m/s}^3)(2.0\ \text{s})^3 = +0.96$ m, and so $\Delta x = +0.96\ \text{m} - (+0.12\ \text{m}) = +0.84$ m. The x component of the average velocity therefore is $(\Delta x)/(\Delta t) = (+0.84\ \text{m})/(1.0\ \text{s}) = +0.84$ m/s. ✔
(*b*) Between $t_i = 1.0$ s and $t_f = 1.1$ s, I have $\Delta x = (+0.12\ \text{m/s}^3)(1.1\ \text{s})^3 - (+0.12\ \text{m}) = +0.039\,72$ m, and $v_{x,\text{av}} = (+0.039\,72\ \text{m})/(0.10\ \text{s}) = +0.3972$ m/s. ✔

Table 2.3 Example 2.12

Δt (s)	Δx (m)	$v_{x,\text{av}} = \Delta x/\Delta t$ (m/s)
1	+0.84	+0.84
0.1	+0.039 72	+0.3972
0.01	+0.003 636 12	+0.363 612
0.001	+0.000 360 360 12	+0.360 360 12
0.0001	+0.000 036 003 600 12	+0.360 036 0012
0.000 01	+0.000 003 600 036 00	+0.360 003 6000
0.000 001	+0.000 000 360 000 36	+0.360 000 3600
0.000 0001	+0.000 000 036 000 00	+0.360 000 0362

(*c*) Table 2.3 shows what happens as I continue to shorten the time interval by factors of 10. As Δt becomes smaller and smaller, $v_{x,\text{av}}$ approaches 0.36 m/s. ✔

❹ **EVALUATE RESULT** If I substitute $c = +0.12\ \text{m/s}^3$ and $t = 1.0$ s into the expression for the x component of the velocity obtained in Example 2.11, I should obtain the same answer as the limiting value for small intervals Δt. Indeed, I find

$$v_x(t) = 3ct^2 = 3(+0.12\ \text{m/s}^3)(1.0\ \text{s})^2 = +0.36\ \text{m/s}.$$



2.19 (*a*) Suppose the position for a certain object is constant. Determine the x component of its velocity by using Eq. 2.22 and by evaluating the limit in Eq. 2.21 of the ratio of the x component of the displacement over a finite time interval. (*b*) Repeat for the case where the x coordinate of the position changes with time according to the expression $x = ct$, where c is a constant. Explain your answer.

Chapter Summary

Average speed and average velocity (Sections 2.2–2.4, 2.6, 2.7)

Concepts The **distance traveled** is the accumulated distance covered by an object along the path of its motion, without regard to direction.

The x component of an object's **displacement** is the change in its x coordinate.

The **average speed** of an object is the distance it travels divided by the time interval it takes to travel that distance.

The x component of an object's **average velocity** is the x component of its displacement divided by the time interval taken to travel that displacement. The magnitude of the average velocity is not necessarily equal to the average speed.

Quantitative Tools The **distance** d between two points x_1 and x_2 is

$$d \equiv |x_1 - x_2|. \quad \text{(one dimension)}. \tag{2.5}$$

The x component of the **displacement** Δx of an object that moves from a point x_i to a point x_f is

$$\Delta x \equiv x_f - x_i. \tag{2.4}$$

The x component of the object's **average velocity** is

$$v_{x,av} \equiv \frac{\Delta x}{\Delta t} = \frac{x_f - x_i}{t_f - t_i}. \tag{2.14}$$

Vectors and scalars (Sections 2.5, 2.6)

Concepts A **scalar** is a physical quantity that is completely specified by a number and a unit of measure. A **vector** is a physical quantity that is completely specified by a number, a unit of measure, and a direction. The number and unit of measure together are called the **magnitude** of the vector.

A **unit vector** has magnitude one and no units.

To **add** two vectors, place the tail of the second vector in the sum at the tip of the first vector; the vector representing the sum runs from the tail of the first vector to the tip of the second. To **subtract** one vector from another, reverse the direction of the vector being subtracted and add the reversed vector to the vector from which you are subtracting.

Quantitative Tools The unit vector $\hat{\imath}$ has magnitude one and points along the x axis in the direction of increasing x.

A vector $\vec{b}$ that points in the x direction and has an x component b_x can be written in **unit vector notation** as

$$\vec{b} \equiv b_x \hat{\imath}. \tag{2.2}$$

Position, displacement, and velocity as vectors (Sections 2.6, 2.7)

Concepts The **position vector,** or **position,** of a point is drawn from the origin of a coordinate system to the point.

The **displacement vector** of an object is the change in its position vector. It is the vector that points from the tip of the original position vector to the tip of the final position vector.

Quantitative Tools The position $\vec{r}$ of a point that has x coordinate x is

$$\vec{r} \equiv x \hat{\imath}. \tag{2.9}$$

The displacement vector $\Delta\vec{r}$ of an object moving along the x axis is

$$\Delta\vec{r} = \vec{r}_f - \vec{r}_i = (x_f - x_i)\hat{\imath}. \tag{2.7, 2.8, 2.10}$$

The average velocity of an object moving along the x axis is

$$\vec{v}_{av} \equiv \frac{\Delta\vec{r}}{\Delta t} = \frac{x_f - x_i}{t_f - t_i}\hat{\imath}. \tag{2.14, 2.15}$$

Additional properties of velocity (Sections 2.8, 2.9)

Concepts When the velocity of an object is constant, its **position-versus-time graph** is a straight line having nonzero slope and its **velocity-versus-time graph** is a horizontal line.

The **instantaneous velocity** of an object is its velocity at any particular instant.

The slope of an object's $x(t)$ curve at a given instant is numerically equal to the x component of the object's velocity at that instant. The area under the object's $v_x(t)$ curve during a time interval is the x component of the object's displacement during that interval.

Quantitative Tools The x component of an object's instantaneous velocity is the derivative of the object's x coordinate with respect to time:

$$v_x \equiv \frac{dx}{dt}. \tag{2.22}$$

Questions and Problems

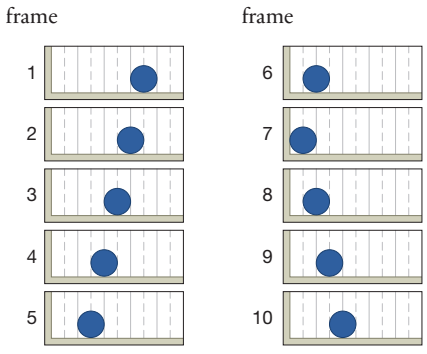
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

2.1 From reality to model

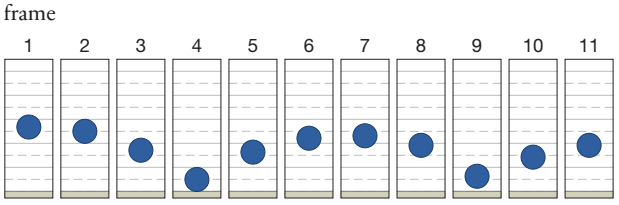
1. What minimum information must be extracted from a video clip of a moving object in order to quantify the object's motion? •
2. The sequence in Figure P2.2 represents a ball rolling into a wall and bouncing off of it. The ball is 10 mm in diameter. Make a graph showing the distance from the leading edge of the ball to the closest part of the wall (using the wall as the origin) as it changes from frame to frame. ••

Figure P2.2



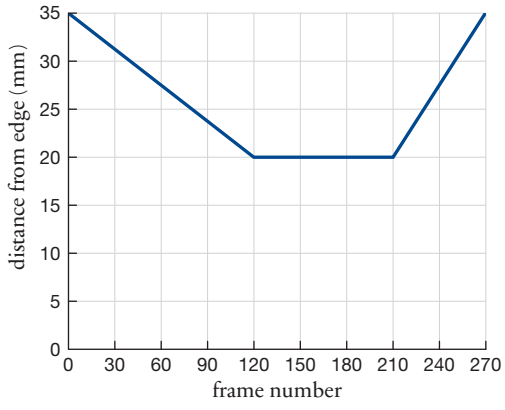
3. The sequence in Figure P2.3 represents a ball that is initially held above the ground. In the first frame the ball is released. In subsequent frames the ball falls, bounces on the ground, rises, and bounces again. The ball is 10 mm in diameter. Make a plot showing the changing height of the ball above the ground. ••

Figure P2.3



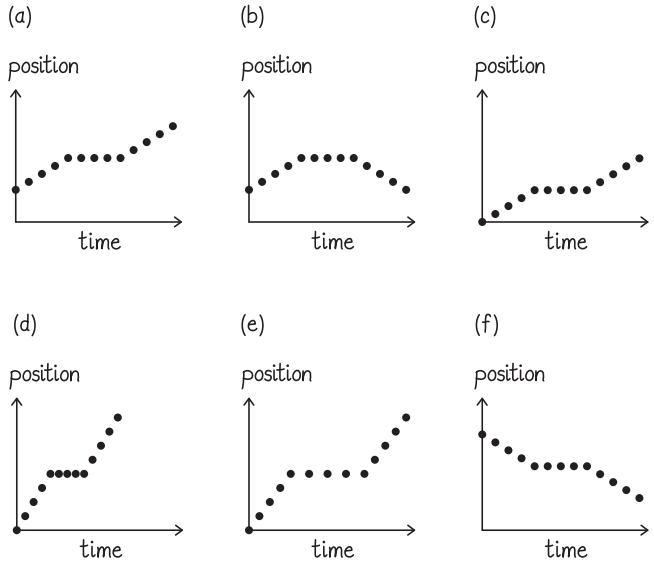
4. Figure P2.4 shows a graph of position versus frame number from a video clip of a moving object. Describe this motion from beginning to end, and state any assumptions you make. ••

Figure P2.4



5. Your class observed several different objects in motion along different lines. Figure P2.5 shows some of the graphs other students made of the events. They have labeled the horizontal axis "time" and the vertical axis "position," but they have not marked any points along those axes and have not specified an origin. Which of the graphs could be describing the same motion? ●●

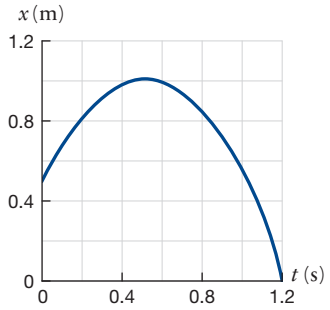
Figure P2.5



2.2 Position and displacement

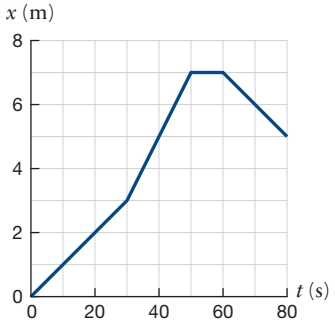
6. A car is initially parked 3.5 m to the left of a fire hydrant. It drives 15.5 m past the hydrant before turning back for 6.0 m and stopping. What is the car's displacement? ●
7. You walk 3.2 km to the supermarket and then back home. What is your distance traveled? What is your displacement? ●
8. You run three times around a 100-m-by-50-m rectangular soccer field and continue for one more length. What is (*a*) the magnitude of your displacement and (*b*) the total distance traveled? ●
9. You are shown a video clip of a dog running in front of a blank wall. The observations are to be plotted using the horizontal and vertical axes labeled with seconds and meters, respectively. How many different graphs of this motion with respect to time could be produced if: (*a*) No additional information is given? (*b*) An origin is given? (*c*) Both the origin and the length of the dog are given? (*d*) An origin, the length of the dog, and the time interval between frames of the video clip are all given? ●●
10. Assume you have a video clip of someone walking from left to right. You draw a position-versus-time graph of the motion and choose your origin to be the left edge of the frame. A friend chooses to take the right edge of the frame as the origin. Both of you choose the positive x direction to be from left to right. (*a*) How would your graphs differ? (*b*) Would the two be equally good representations of reality? ●●
11. Figure P2.11 shows the position measured in meters plotted against time measured in seconds for an object moving in one dimension. For motion between $t = 0$ and $t = 1.2$ s, (*a*) describe the motion of this object. (*b*) What is the distance traveled by this object? (*c*) What is the displacement of the object? ●●

Figure P2.11



12. In the graph in Figure P2.12, determine (*a*) the displacement of the object and (*b*) the distance the object traveled. ●●

Figure P2.12

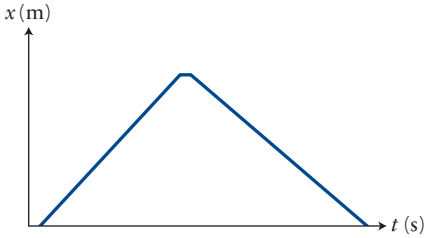


13. An ant walks along a tight thread that is oriented north to south and can be taken to coincide with the y axis. The ant walks 7 cm south, then turns and walks 2 cm north. It continues another 10 cm north, then 3 cm south, then 12 cm north, then 6 cm south. What is the y component of the ant's displacement in cm? Take the origin to be the ant's starting point. ●●
14. In the foreground of a side-view picture of a table, the legs touch the floor at a point that is 12 mm from the bottom of the picture, and the tabletop is 65 mm from the bottom of the picture. One end of the tabletop is 14 mm from the left edge of the picture, and the other end is 99 mm from the left edge. Estimate the actual length of the tabletop. ●●●

2.3 Representing motion

15. Figure P2.15 shows the position of a swimmer in a race as a function of time. Describe this motion. ●

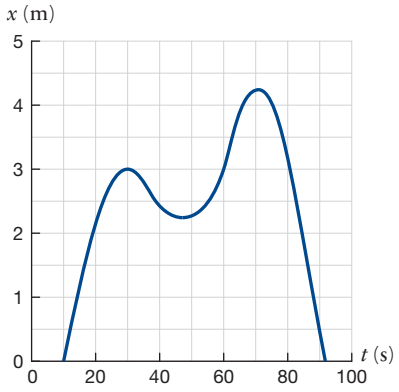
Figure P2.15



16. In the morning, a hiker at the bottom of a mountain heads up the trail toward the top. At the same instant, another hiker at the top of the mountain heads down the same trail toward the bottom. Each hiker arrives at her destination by the end of the day. Explain why, no matter what happens to either hiker along the way (lunch breaks, turnarounds to fetch dropped pocket knives, or changes of pace), they meet on the trail sometime during the day. ●

17. Will interpolation between known data points always give an accurate continuous path? If so, explain why. If not, give a counterexample. •
18. Figure P2.18 shows the motion of an object as a function of time. How long did it take the object to get from the position $x = 2.0$ m to the position $x = 3.0$ m? Is there only one correct answer? ••

Figure P2.18



19. Figure P2.19 is based on a multiple-flash photographic sequence of a ball rolling from left to right on a soft surface. Plot the ball's position as a function of time. ••

Figure P2.19



20. A remotely controlled helicopter is heading vertically upward when you attempt to land it. Its height is represented by $y(t) = p + qt - rt^2$, where $p = 20.0$ m, $q = 5.0$ m/s, and $r = 5.0$ m/s^2. (*a*) Draw a graph of the motion from $t = 0$ to $t = 2$ s. (*b*) What was the initial position of the helicopter? (*c*) What is the maximum height that it reaches before descending? (*d*) At what time does the helicopter land? ••
21. After a rocket is launched at $t = 0$, its position is given by $x(t) = qt^3$, where q is some positive constant. (*a*) Sketch a graph of the rocket's position as a function of time. (*b*) What is the rocket's displacement during the interval from $t = T$ to $t = 3T$? ••
22. Consider the position function $x(t) = p + qt + rt^2$ for a moving object, with $p = +3.0$ m, $q = +2.0$ m/s, and $r = -5.0$ m/s^2. (*a*) What is the value of $x(t)$ at $t = 0$? (*b*) At what value of t does $x(t)$ have its maximum value? (*c*) What is the value of $x(t)$ at this instant? (*d*) Draw a position-versus-time graph. (*e*) Describe the behavior of an object that is represented by this function. (*f*) How far has the object traveled in these intervals: From $t = 0$ to $t = 0.50$ s? From $t = 0$ to $t = 1.0$ s? From $t = 0.50$ s to $t = 1.0$ s? ••
23. The motion of some object is described by the equation $x(t) = at - b \sin(ct)$, where $a = 1.0$ m/s, $b = 2.0$ m, and $c = 4\pi$ s^{-1}. This motion is being observed by four students, and measurements of the position of this object begin at time $t = 0$. Students A, B, C, and D measure every 1.0 s, 0.5 s, 0.25 s, and 0.1 s, respectively. Draw four graphs of the measurements made by the four different students. Include at least ten data points in each graph. Do the four students agree on the type of motion that is taking place? •••

2.4 Average speed and average velocity

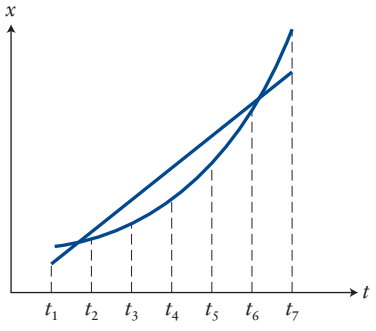
24. In the Midwest, you sometimes see large marks painted on the highway shoulder. How can police patrols flying overhead use these marks to check for speeders? •
25. Calculate the average speed for the runners in the following races: (*a*) 100 m in 9.84 s, (*b*) 200 m in 19.32 s, (*c*) 400 m in 43.29 s, (*d*) 1500 m in 3 min, 27.37 s, (*e*) 10 km in 26 min, 38.08 s, (*f*) marathon (26 mi, 385 yd) in 2 h, 6 min, 50.0 s. •
26. (*a*) Can two cars traveling in opposite directions on a highway have the same speed? (*b*) Can they have the same velocity? •
27. Figure P2.27 is based on a multiple-flash photographic sequence, taken at equal time intervals, of a ball rolling on a smooth surface from right to left. (*a*) Argue that the ball moves with at least two different speeds during the motion represented in the sequence. (*b*) During which part of the motion does the ball have the larger speed? •

Figure P2.27



28. Can the average speed of an object moving in one direction ever be larger than the object's maximum speed? ••
29. Figure P2.29 shows position as a function of time for two cars traveling along the same highway. (*a*) At what instant(s) are the cars next to each other? (*b*) At what instant(s) are they traveling at the same speed? ••

Figure P2.29



30. Figure P2.30 is based on two multiple-flash photographic sequences of a hockey puck sliding on ice. Sequence *a* was shot at 30 flashes per second, and sequence *b* was shot at 20 flashes per second. Compare the speeds of the puck in the two cases. ••

Figure P2.30

(*a*)



(*b*)



31. A cyclist takes 10 min to ride from point A to point B and then another 10 min to continue on from point B to point C, all along a straight line. If you know that the average speed on the ride from A to B was faster than the average speed on the ride from B to C, what, if anything, can you conclude about the position of point B relative to points A and C? ••
32. In a road rally race, you are told to drive half the trip at 25 m/s and half the trip at 35 m/s. It's not clear from the directions whether this means to drive half the *time* at each speed or drive half the *distance* at each speed. Which would yield the shorter travel time for the entire trip? ••
33. A bicycle racer rides from a starting marker to a turnaround marker at 10.0 m/s. She then rides back along the same route from the turnaround marker to the starting marker at 16.0 m/s. (*a*) What is her average speed for the whole race? (*b*) A friend of yours gets an answer of 13 m/s. What is most likely the source of his error? ••
34. A taxi driver leaves the airport to drop a passenger at his hotel, making the following trips: she drives 7.2 km in 10 min, stops and waits for the passenger at a convenience store for 7 min, drives another 5 km to the hotel in 10 min, where she waits 3 min for the passenger to get off. She then returns to the airport via a different route that is 10 km long in 6 min. (*a*) What is her average speed for the whole trip? (*b*) What is her average velocity? ••
35. You are going to visit your grandparents, who live 500 km away. As you drive on the freeway, your speed is a constant 100 km/h. Half an hour after you leave home, your brother discovers that you forgot your wallet. He jumps into his car and speeds after you. The two of you arrive at your grandparents' house at the same instant. What is your brother's average speed for the trip? ••
36. You and your brother both leave your house at the same instant and drive in separate cars along a straight highway to a nearby lake. After 10 min, you are both 3.0 km from your house. You are now driving at 100 km/h, and you continue at this constant speed; your brother is going faster than this. After another 20 min, your brother arrives at the lake, but you are still 5.0 km away. (*a*) How far is it from your house to the lake? (*b*) What is your brother's average speed for the trip? (*c*) Your brother beats you to the lake by what time interval? •••

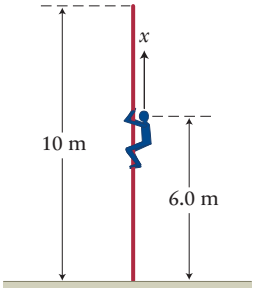
2.5 Scalars and vectors

37. You are standing on a sidewalk that runs east-west. Consider these instructions I might give you: (1) Walk 15 steps along the sidewalk and stop. (2) Walk 15 steps westward on the sidewalk and stop. Do my instructions unambiguously determine your final location in each case? •
38. What is the x component of (*a*) $(-1.2\text{ m})\hat{\imath}$, (*b*) $(3.2\text{ m})\hat{\imath}$, and (*c*) $(-1.2\text{ m/s})\hat{\imath}$? •
39. What is the magnitude of (*a*) $(+3\text{ m})\hat{\imath}$, (*b*) $(+3\text{ m/s})\hat{\imath}$, and (*c*) $(-3\text{ m/s})\hat{\imath}$? •
40. Vectors $\vec{A}$ and $\vec{B}$ each have a magnitude of 5 m and point to the left. Vector $\vec{A}$ begins at the origin, while vector $\vec{B}$ begins at a location 8 m to the right of the origin. (*a*) If the positive x axis is pointed to the right, what is the x component of each vector? (*b*) If the positive x axis is pointed to the left, what is the x component of each vector? ••
41. Vector $\vec{A}$ points to the right, as does the positive x axis. (*a*) Express this vector in unit vector notation. (*b*) Now flip $\vec{A}$ to the opposite direction. Express it in unit vector notation. (*c*) Keep the vector in its new direction and flip the x axis so that the positive direction is to the left. Express the vector in unit vector notation. ••

2.6 Position and displacement vectors

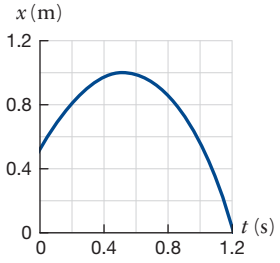
42. Consider three vectors along the x axis, $\vec{A}$, $\vec{B}$, and $\vec{C}$. Let the x components of the three vectors be -2 m, $+3$ m, and -5 m. What are (*a*) $\vec{D} = \vec{A} + \vec{B}$ and (*b*) $\vec{E} = \vec{D} - \vec{C}$? •
43. You stop to rest while climbing a vertical 10-m pole. With the origin at the level of your head and with the positive x direction upward, as shown in Figure P2.43, what are (*a*) the x coordinate of the pole's tip and (*b*) the tip's position vector? •

Figure P2.43



44. The height x above the ground of a vertically launched projectile is given by $x(t) = pt - qt^2$, with $p = 42$ m/s and $q = 4.9\text{ m/s}^2$. (*a*) At what instant is the projectile at a height of 20 m? (*b*) What is the meaning of the two solutions obtained in part *a*? (*c*) Sketch a graph of the x component of the projectile's velocity as a function of time. ••
45. Figure P2.45 shows the x coordinate as a function of time for a moving object. What is the total distance traveled by the object (*a*) between $t = 0$ and $t = 0.45$ s and (*b*) between $t = 0$ and $t = 1.2$ s? What is the displacement of the object (*c*) between $t = 0$ and $t = 0.45$ s and (*d*) between $t = 0$ and $t = 1.2$ s? Which way is the object moving at (*e*) $t = 0.2$ s, (*f*) $t = 0.45$ s, and (*g*) $t = 1.0$ s? ••

Figure P2.45

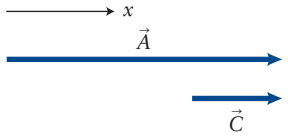


46. Arrange three displacement vectors, of magnitudes 2 m, 5 m, and 7 m, so that their sum is (*a*) $(+10\text{ m})\hat{\imath}$, (*b*) $(-4\text{ m})\hat{\imath}$, and (*c*) 0. ••
47. The direction of vector $\vec{A}$ is opposite the direction of the unit vector $\hat{\imath}$. Vector $\vec{B}$ has half the magnitude of $\vec{A}$, and $\vec{A} - \vec{B}$ is a vector of magnitude $\frac{3}{2}A$. Express $\vec{B}$ in terms of $\vec{A}$. ••
48. You have to deliver some 5.0-kg packages from your home to two locations. You drive for 2.0 h at 25 mi/h due east (call this segment 1 of your trip), then turn around and drive due west for 30 min at 20 mi/h (segment 2). Use a

coordinate system with the positive x axis aimed toward the east and the origin at your home. (*a*) What is your position vector at the instant you reach the end of segment 1? (*b*) What is your position vector at the instant you reach the end of segment 2? (*c*) Calculate your displacement during segment 2. (*d*) Calculate your displacement for the whole trip. (*e*) What is the distance traveled? (*f*) Draw to scale the position vectors you found in parts *a* and *b*. (*g*) Use vector addition to determine the displacement vectors asked for in parts *c* and *d*. Do your results in part *g* agree with the values you calculated in parts *c* and *d*? ●●

49. (*a*) In Figure P2.49, what vector must you add to $\vec{A}$ to get $\vec{C}$? (*b*) What vector must you subtract from $\vec{A}$ to get $\vec{C}$? Sketch your answers on a copy of the figure to confirm your results. ●●

Figure P2.49



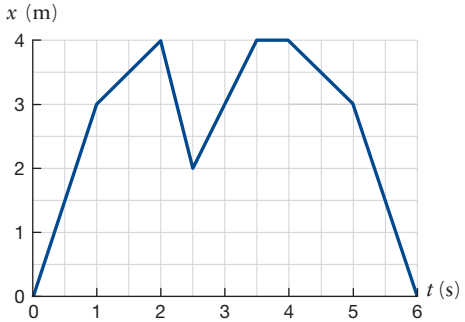
2.7 Velocity as a vector

50. After driving due north at 72 km/hr for 2 hours, you stop for 20 min and then head back south at 108 km/hr to reach your starting position. What is your (*a*) average velocity and (*b*) average speed at the end of the trip? ●

51. For the motion represented in Figure P2.45, calculate (*a*) the object's average velocity between $t = 0$ and $t = 1.2$ s, and (*b*) its average speed during this same time interval. (*c*) Why is the answer to part *a* different from the answer to part *b*? ●●

52. Figure P2.52 is the position-versus-time graph for a moving object. What is the object's average velocity (*a*) between $t = 0$ and $t = 1.0$ s, (*b*) between $t = 0$ and $t = 4.0$ s, and (*c*) between $t = 3.0$ s and $t = 6.0$ s? (*d*) What is its average speed between $t = 3.0$ s and $t = 6.0$ s? (*e*) Draw to scale the velocity vector for the intervals $t = 3.0$ s to $t = 3.5$ s, $t = 3.5$ s to $t = 4.0$ s, $t = 4.0$ s to $t = 5.0$ s, and $t = 5.0$ s to $t = 6.0$ s. (*f*) What is the sum of all the vectors you drew in part *e*? ●●

Figure P2.52



53. You normally drive a 12-h trip at an average speed of 100 km/h. Today you are in a hurry. During the first two-thirds of the distance, you drive at 108 km/h. If the trip still takes 12 h, what is your average speed in the last third of the distance? ●●●

2.8 Motion at constant velocity

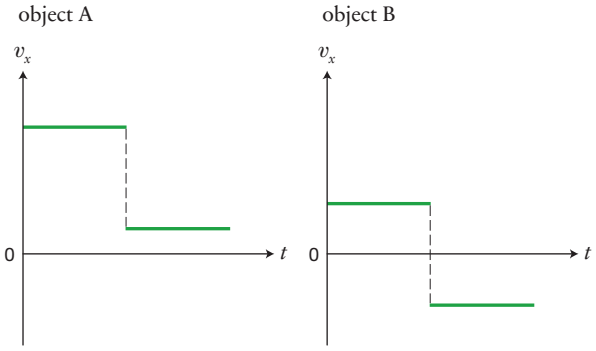
54. A cart starts at position $x = -2.073$ m and travels along the x axis with a constant x component of velocity of -4.02 m/s. What is the position of the cart after 0.103 s? ●

55. A bug on a windowsill walks at 10 mm/s from left to right for 120 mm, slows to 6.0 mm/s and continues rightward for another 3.0 s, stops for 4.0 s, and then walks back to its starting position at 8.0 mm/s. Draw a graph (*a*) of the x coordinate of the bug's displacement as a function of time and (*b*) of the x component of its velocity as a function of time. ●

56. A freight train leaves city A and heads for city B, which is 44 km away, at 12 km/hr. At the same time, a passenger train leaves city B and heads for A. The two trains barely avoid a collision at a station exactly 12 km from city A. At what constant velocity was the passenger train traveling? ●●

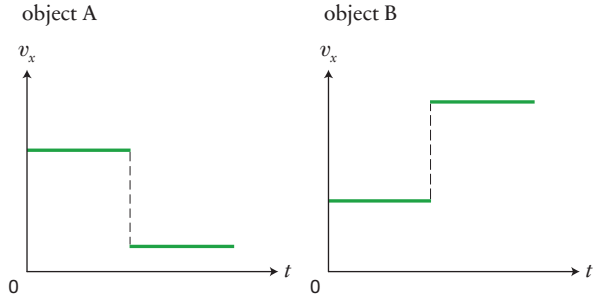
57. Figure P2.57 shows the velocity-versus-time graphs for objects A and B moving along an x axis. Which object has the greater displacement over the time interval shown in the graph? ●●

Figure P2.57



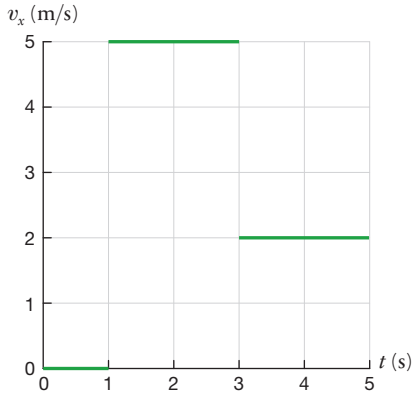
58. Figure P2.58 shows the x component of the velocity as a function of time for objects A and B. Which object has the greater displacement over the time interval shown in the graph? ●●

Figure P2.58



59. An object moving along an x axis starts out at $x = -10$ m. Using its velocity-versus-time graph in Figure P2.59, draw a graph of the object's x coordinate as a function of time. ●●

Figure P2.59



60. You and a friend ride bicycles to school. Both of you start at the same instant from your house, you riding at 10 m/s and your friend riding at 15 m/s. During the trip your friend has a flat tire that takes him 12 min to fix. He then continues the trip at the same speed of 15 m/s. If the distance to school is 15 km, which of you gets to school first? ●●

61. You are going on a bicycle ride with a friend. You start 3.0 min ahead of her from her carport, pedal at 5.0 m/s for 10 min, then stop and chat with a neighbor for 5.0 min. As you chat, your friend pedals by, oblivious to you. You notice that she's forgotten to bring the lunch basket you plan to share, and you immediately start to pedal back to her house to get it, now moving at 10 m/s. When you are halfway back, your friend catches up with you, having herself remembered about the basket. (*a*) What was the magnitude of her average velocity for the trip up to that instant? (*b*) What was the magnitude of her average velocity between the instant she passed you and the neighbor and the instant she caught up with you? ●●●

62. A policeman starts giving chase 60 seconds after a stolen car zooms by at 108 km/hr. At what minimum speed should they drive if they have to catch up with the driver before they manage to get onto an expressway 60 km away? ●●●

63. You are jogging eastward at an average speed of 2.0 m/s. Once you are 2.0 km from your home, you turn around and begin jogging westward, back to your house. At one point on your return trip, you look to the north and see a friend. You run northward at a greater speed of 3.0 m/s to catch your friend. After jogging next to each other for a time (at 3.0 m/s), you turn around (southward) and return at 3.0 m/s to your original path, and resume jogging at 2.0 m/s westward. If you return home 40 min after you left, how far north did you run while catching and running with your friend? ●●●

2.9 Instantaneous velocity

64. Which of these quantities depend on the choice of origin in a coordinate system: position, displacement, speed, average velocity, instantaneous velocity? ●

65. An object's position as a function of time is given by $x(t) = -2.0t^2 + 1.0\,t^3$. Calculate (*a*) the object's position and velocity at $t = 2$ s. (*b*) The average velocity between $t = 0$ and $t = 2$ s. ●

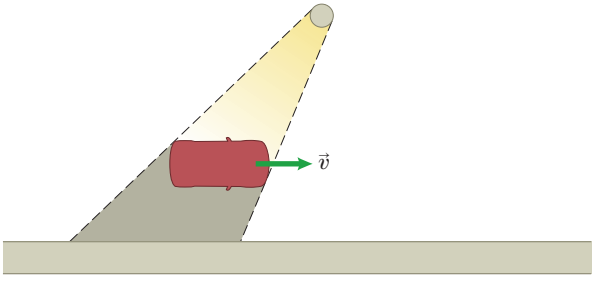
66. A mouse runs along a baseboard in your house. The mouse's position as a function of time is given by $x(t) = pt^2 + qt$, with $p = 0.40$ m/s^2 and $q = -1.20$ m/s. Determine the mouse's average velocity and average speed (*a*) between $t = 0$ and $t = 1.0$ s and (*b*) between $t = 1.0$ s and $t = 4.0$ s. ●●

67. Car A is spotted passing car B just east of Westerville at exactly 2:00 p.m. The same two cars are then spotted next to each other just west of Easterville at exactly 3:00 p.m. If car B had a constant velocity of 30 m/s eastward for the whole trip, defend the proposition that car A had to have a velocity of 30 m/s eastward at some instant during that hour. ●●

68. The motion of an electron is given by $x(t) = pt^3 + qt^2 + r$, with $p = -2.0$ m/s^3, $q = +1.0$ m/s^2, and $r = +9.0$ m. Determine its velocity at (*a*) $t = 0$, (*b*) $t = 1.0$ s, (*c*) $t = 2.0$ s, and (*d*) $t = 3.0$ s. ●●

69. The car in Figure P2.69 passes a bright streetlight at constant speed, casting a shadow on a wall on the other side of the street. For simplicity, assume that car and light are at the same height. (*a*) Which is greater: the average speed of the car or the average speed of the shadow's leading edge? (*b*) Is there any instant at which the speed of the shadow's leading edge is the same as the speed of the car? If so, what is the car's position relative to the streetlight at that instant? If not, why not? ●●

Figure P2.69



70. The position of a 6.0-kg shopping cart rolling down a ramp is given by $x(t) = p + qt^2$, with $p = +1.50$ m and $q = +2.00$ m/s^2. What is the x component of the cart's average velocity (*a*) between $t = 2.00$ s and $t = 3.00$ s, (*b*) between $t = 2.00$ s and $t = 2.10$ s, and (*c*) between $t = 2.00$ s and $t = 2.01$ s? (*d*) Compute the limit of the average velocity between $t_i = 2.00$ s and $t_f = 2.00\text{ s} + \Delta t$ as Δt approaches zero. (*e*) Show that your result agrees with what is expected by taking the time derivative of position. ●●●

Additional Problems

71. You leave Fort Worth, Texas, at 2:38 p.m. and arrive in Dallas at 3:23 p.m., covering a distance of 58 km. What is your average speed (*a*) in meters per second and (*b*) in miles per hour? •
72. In a footrace between two runners, is it possible for the second-place finisher to have a greater speed at the finish line than the winner? •
73. You wish to describe the position of the base of the pole in Figure P2.43 using the indicated coordinate system. What are (*a*) the position coordinate of the base, (*b*) the position vector of the base, and (*c*) the magnitude of that vector? •
74. Runners P, Q, and R run a 5-km race in 15, 20, and 25 min, respectively, each at a constant speed. When runner Q crosses the 1-km mark, what is the distance, to the nearest meter, between runners P and R? Assume that times and distances are correct to at least four significant digits. ••
75. Runners A, B, and C run a 100-m race, each at a constant speed. Runner A takes first place, beating runner B by 10 m. Runner B takes second place, beating runner C by 10 m. By what time interval does runner A beat runner C? (Express your answer in terms of runner A's total time.) ••
76. At $t = 0$, car A passes a milepost at constant speed v_A. Car B passes the same milepost at constant speed $v_B > v_A$ after a time interval Δt has elapsed. (*a*) In terms of v_A, v_B, and Δt, at what instant does car B catch up to car A? (*b*) How far past the milepost does this happen? ••
77. You and your friend are running at a long racetrack. You pass the starting line while running at a constant 4.0 m/s. Fifteen seconds later, your friend passes the starting line while running at 6.0 m/s in the same direction, and at the same instant you increase your speed to 5.0 m/s. (*a*) How long does it take for your friend to catch up with you? (*b*) How far are the two of you from the starting line when your friend catches up with you? ••
78. You drive an old car on a straight, level highway at 45 mi/h for 10 mi, and then the car stalls. You leave the car and, continuing in the direction in which you were driving, walk to a friend's house 2.0 mi away, arriving 40 min after you begin walking. What is your average speed during the whole trip? ••
79. An object's position is represented by $x(t) = 2.0\,t^{\frac{3}{2}}$. Compare the instantaneous velocity of the object at $t = 4$ s to its average velocity between $t = 0$ and $t = 4$ s. ••
80. The following equations give the x component of the position for four objects as functions of time:

    object 1: $x(t) = a$, where $a = 5$ m

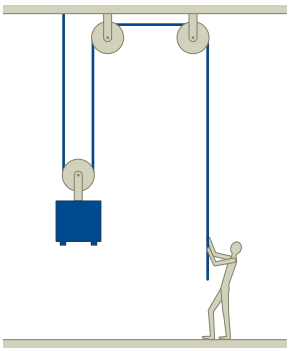
    object 2: $x(t) = bt + c$, where $b = +4$ m/s and $c = -1$ m

    object 3: $x(t) = et^2 + ft$, where $e = +5$ m/s^2 and $f = -9$ m/s

    object 4: $x(t) = gt^2 + h$, where $g = -3$ m/s^2 and $h = +12$ m.

    (*a*) Which objects have a velocity that changes with time? (*b*) Which object is at the origin at the earliest instant, and what is that instant? (*c*) What is that object's velocity 1 s after the instant you calculated in part *b*? ••
81. A furniture mover is lifting a small safe by pulling on a rope threaded through the pulley system shown in Figure P2.81. (*a*) What is the ratio of the vertical distance the safe moves to the length of the rope pulled by the mover? (*b*) What is the ratio of the speed of the safe to the speed of the segment of rope pulled by the mover? ••

Figure P2.81



82. Two steamrollers begin 100 m apart and head toward each other, each at a constant speed of 1.00 m/s. At the same instant, a fly that travels at a constant speed of 2.20 m/s starts from the front roller of the southbound steamroller and flies to the front roller of the northbound one, then turns around and flies to the front roller of the southbound once again, and continues in this way until it is crushed between the steamrollers in a collision. What distance does the fly travel? ••
83. Consider a 2.0-kg object that moves along the x axis according to the expression $x(t) = ct^3$, where $c = +0.120$ m/s^3. (*a*) Determine the x component of the object's average velocity during the interval from $t_i = 0.500$ s to $t_f = 1.50$ s. (*b*) Repeat for the interval from $t_i = 0.950$ s to $t_f = 1.05$ s. (*c*) Show that your results approach the x component of the velocity at $t = 1.00$ s if you continue to reduce the interval by factors of 10. Use all significant digits provided by your calculator at each step. ••
84. The position of a yo-yo as a function of time is given by $x(t) = A\cos(pt + q)$, where $A = 0.60$ m, $p = \frac{1}{2}\pi$ s^{-1}, and $q = \frac{1}{2}\pi$. (*a*) Plot this function at 17 equally spaced instants from $t = 0$ to $t = 8.0$ s. (*b*) At what instants is the velocity zero? (*c*) Plot the x component of the velocity as a function of time over the time interval from $t = 0$ to $t = 8$ s. •••
85. Zeno, a Greek philosopher and mathematician, was famous for his paradoxes, one of which can be paraphrased as follows: A runner has a race of length d to run. After an elapsed time interval Δt after the start, he is a distance $\frac{1}{2}d$ from the finish line. After an additional time interval of $\frac{1}{2}\Delta t$ (elapsed time is $\Delta t + \frac{1}{2}\Delta t$), he has traveled a distance of $\frac{3}{4}d$. He then is a distance $\frac{1}{4}d$ from the finish line. At elapsed time $\Delta t + \frac{1}{2}\Delta t + \frac{1}{4}\Delta t$, he is a distance $\frac{1}{8}d$ from the finish line, and so on. This description of the motion is an infinite series of smaller and smaller displacements. The paradox is that this analysis suggests that the runner will never make it to the finish line, even though we know he does. (*a*) What is the runner's speed? (*b*) What time interval is needed for the trip? (*c*) Describe the resolution to this seeming paradox. •••

86. Four traffic lights on a stretch of road are spaced 300 m apart. There is a 10-s lag time between successive green lights: The second light turns green 10 s after the first light turns green, the third light turns green 10 s after the second light does (and thus 20 s after the first light turns green), and so on. Each light stays green for 15 s. Having watched this sequence for some time, you understand that there should be one best constant speed to drive to make all four green lights. You do not want to be delayed, but you are a cautious driver and want to drive as slowly as you can but still make it through all the lights on a single cycle. ●●● CR

87. Hare and Tortoise of Aesop's fable fame are having a rematch, a mile-long race. Hare has planned more carefully this time. Five minutes into the race, he figures he can take a 40-min nap and still win easily because it takes him only 10 min to cover 1 mi and because he thinks that Tortoise's top speed is 1.0 mi/h. However, Tortoise has been working out, and as he passes napping Hare, he suddenly increases his speed to $\frac{5}{3}$ mi/h.When Hare wakes up, he sees Tortoise way ahead and runs at full speed to try to catch up. ●●● CR

88. Two friends are running a 1000-m race. Jim gets a head start on Jack by starting earlier. How much earlier should Jim start to win the race if he can only run at 6.8 m/s while Jack can easily run at 7 m/s? ●●●

89. You are on planet Dither, whose inhabitants often change their minds on how to choose a reference axis. At time $t = 0$, while standing 2.0 m to the left of the origin of a reference axis for which the positive x direction points to the left, you launch a toy car that moves with increasing speed to the right. The car's x coordinate (x component of the position) is given by $x(t) = p + qt + rt^2$, where $p = +2.0$ m, $q = -3.0$ m/s, and $r = -4.0$ m/s^2. (*a*) What is the equation of the car's x coordinate if the reference axis is flipped so that the positive direction points to the right? (*b*) What is the equation of the car's x coordinate if the axis is kept in its original orientation (leftward is the positive direction) but the origin is moved to where you are standing? (*c*) Draw a position-versus-time graph for each of the three combinations of axis direction and origin location. (*d*) What is the x component of the car's displacement at $t = 4.0$ s in each of these three axis/origin combinations? (*e*) What is the x component of the car's velocity at this instant in each combination? (*f*) Explain why there is no physical difference in the answers to parts *d* and *e* even though some of the mathematical expressions you get look different from one another. ●●●

90. Your dream job as a 12-year-old was to sit at a computer at NASA mission control and guide the motion of the Mars rover across the Martian surface 2.0×10^8 km away. Communication signals to steer the rover travel at the speed of light between Earth and Mars, and the rover's top speed is 2.0 m/min. In your dream, your job was to make sure, based on television images beamed back from the craft, that the rover does not run off a Martian cliff. The shift supervisor recommended that your first task be to determine how far ahead in the rover's field of view you have to watch out for obstacles in order to avoid a disaster. ●●● CR

91. You are a driver for Ace Mining Company. The boss insists that, every hour on the hour, a loaded truck leaves the mine at 90.0 km/h, carrying ore to a mill 630 km away. She also insists that an empty truck traveling at 105 km/h leaves the mill every hour on the half hour to return to the mine. One day you begin to count the empty trucks that you encounter, per hour, on your trip to the mill. The next day you decide that loaded trucks are more important, counted on your return trip. ●●● CR



3

Acceleration

CONCEPTS

3.1 Changes in velocity
3.2 Acceleration due to gravity
3.3 Projectile motion
3.4 Motion diagrams

QUANTITATIVE TOOLS

3.5 Motion with constant acceleration
3.6 Free-fall equations
3.7 Inclined planes
3.8 Instantaneous acceleration

In Section 2.8 we considered objects moving at constant velocity—unchanging speed and unchanging direction. Most motions, however, are not at constant velocity because things speed up and slow down and change direction all the time. Any change in velocity is caused by *acceleration,* the subject of this chapter.

3.1 Changes in velocity

Consider a car at rest at a traffic light. The light turns green, and the driver puts the car in motion. As long as the car's velocity is changing, the car is *accelerating:*

If an object's velocity is changing, the object is accelerating.

An object's velocity can change quickly or slowly: A sports car takes off much more quickly than a heavily loaded truck. Starting from rest, a sports car can reach a speed of 30 m/s in as little as 5 s; a truck might take minutes to reach the same speed. The difference between the two motions is the *rate* at which the speeds change. For example, if a sports car's speed increases from 0 to 10 m/s in 2.5 s, the rate at which the speed changes is $(10\ \text{m/s} - 0)/(2.5\ \text{s}) = 4.0\ \text{m/s}^2$. This rate, which is typically greater for a sports car than for a truck, is the magnitude of a quantity we call *acceleration*. Like velocity, acceleration is a vector, and so we need not only a magnitude but also a direction to specify it completely. We represent acceleration by the symbol $\vec{a}$.

For now, let's return to the example of the sports car and choose the x axis along its direction of travel. The rate at which the x component of the car's velocity changes is called the x component of its **average acceleration:**

The x component of the average acceleration of an object is the change in the x component of the velocity divided by the time interval during which this change took place.

I deliberately use the word *average* here because nothing prevents the x component of the car's velocity from increasing at different rates. For example, the x component of its velocity could increase from 0 to +5.0 m/s in 1.0 s, and then from +5.0 m/s to +10 m/s in 1.5 s.

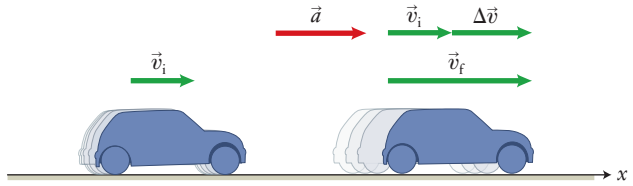
The SI unit of acceleration, m/s^2 (meters per second squared), is a shorthand way of writing (m/s)/s (meters per second, per second). If something accelerates at $+1\ \text{m/s}^2$, the x component of its velocity increases by +1 m/s each second. One of the units of time in the s^2 comes from the change in velocity; the other comes from the time interval during which this change occurs.



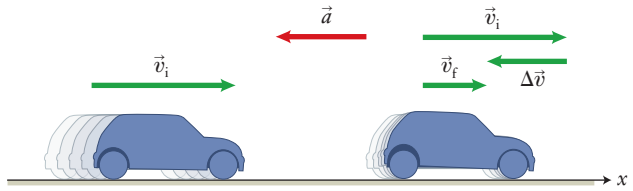
3.1 The x component of a car's velocity increases from 0 to +5.0 m/s in 1.0 s, and then from +5.0 m/s to +10 m/s in the next 2.0 s. What is the x component of its average acceleration (*a*) during the first second, (*b*) during the last 2 seconds, and (*c*) during the entire 3.0-s interval?

Figure 3.1 The direction of the car's acceleration depends on the direction of the x axis and on whether the car is speeding up or slowing down.

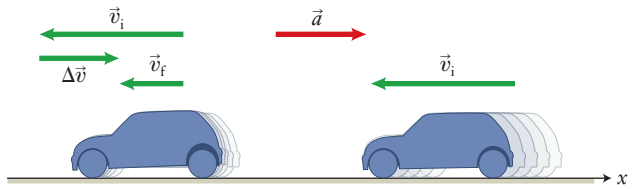
(*a*) Car speeds up in positive x direction



(*b*) Car slows down in positive x direction



(*c*) Car slows down in negative x direction



To see how the direction of acceleration relates to the motion, consider **Figure 3.1**. In Figure 3.1*a* a car traveling in the positive x direction (which means that its velocity points in the positive x direction) speeds up. The magnitude of the car's final velocity $\vec{v}_f$ is greater than the magnitude of the car's initial velocity $\vec{v}_i$, and so the change in the car's velocity $\Delta\vec{v} \equiv \vec{v}_f - \vec{v}_i$, which points from the tip of $\vec{v}_i$ to the tip of $\vec{v}_f$ (see Figure 2.27*a*), is in the positive x direction. By definition the acceleration $\vec{a}$ always points in the same direction as $\Delta\vec{v}$, and so here $\vec{a}$ points in the positive x direction. Consequently, the x component of the car's acceleration is positive: $a_x > 0$.

The car in Figure 3.1*b* also travels in the positive x direction, but this car slows down, and so $\Delta\vec{v}$ now points in the negative x direction—opposite the velocity vector. The acceleration vector, again pointing in the same direction as $\Delta\vec{v}$, also points in the negative x direction, giving a negative x component of the car's acceleration: $a_x < 0$.

Because its velocity is decreasing, you may be tempted to say that the car in Figure 3.1*b* is "decelerating"—an everyday expression used to indicate that something is slowing down. In physics, however, the word *deceleration* is not necessary and therefore best avoided; *acceleration* refers to *any* change in velocity, regardless of whether the object is speeding up or slowing down.

CONCEPTS

Figure 3.1*c* shows a surprising fact: A positive a_x does not have to mean speeding up. As you can see from the figure, the car slows down while traveling in the negative x direction. Because of the direction of travel, the car's velocity $\vec{v}$ now points in the negative x direction. The vector $\Delta\vec{v}$, however, points in the positive x direction (still from the tip of $\vec{v}_i$ to the tip of $\vec{v}_f$). The acceleration therefore also points in the positive x direction, and so the x component of the car's acceleration is *positive* even though the car slows down!

Whenever an object's velocity and acceleration vectors point in the same direction, the object speeds up. When they point in opposite directions, the object slows down.

3.2 The x component of the velocity of a car changes from -10 m/s to -2.0 m/s in 10 s. (*a*) Is the car traveling in the positive or negative x direction? (*b*) Does $\Delta\vec{v}$ point in the positive or negative x direction? (*c*) Is the x component of the car's acceleration positive or negative? (*d*) Is the car speeding up or slowing down?

When working with objects traveling in the negative x direction, don't confuse "positive x component of acceleration" with "speeding up." The x component of velocity in Figure 3.1*c*, for example, is increasing (it gets less negative), and so a_x is positive even though the car is slowing down. Something speeds up when the *magnitude* of its velocity (which, remember from Chapter 2, is called its *speed*) increases, regardless of which way the object is moving. Whether or not the x component of an object's acceleration is positive, however, depends on the direction in which the object is moving.

As mentioned in Section 2.8, the $x(t)$ curve for an object moving at constant velocity is a straight line (see Figure 2.29*a*). When the velocity is not constant, the $x(t)$ curve is not straight. In **Figure 3.2*a***, the x component of displacement Δx_2 during a certain time interval Δt beginning at t_2 is larger than the x component of displacement Δx_1 during the same time interval beginning at t_1. This means that the x component of the object's velocity is increasing. In Figure 3.2*b*, where the

Figure 3.2 Position-versus-time graphs for objects accelerating along the positive x axis.

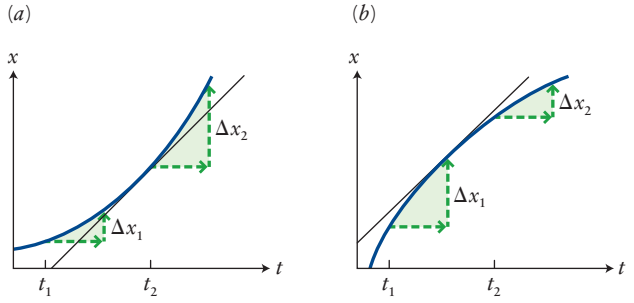
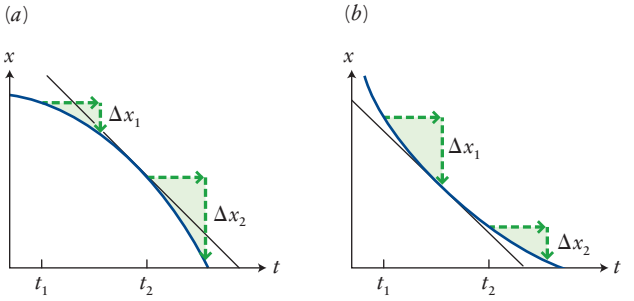


Figure 3.3 Position-versus-time graphs for objects accelerating along the positive x axis.



$x(t)$ curve bends downward, the x component of displacement Δx_2 beginning at t_2 is smaller than the x component of displacement Δx_1 beginning at t_1, telling us that the x component of the velocity is decreasing.

3.3 (*a*) In **Figure 3.3**, are the x components of the velocity represented by the $x(t)$ curves positive or negative? (*b*) Are the speeds increasing or decreasing? (*c*) Are the x components of $\Delta\vec{v}$ positive or negative? (*d*) Are the x components of the acceleration positive or negative?

If you answered Checkpoint 3.3 correctly, you are to be congratulated. Getting this right the first time around is not easy.

An interesting point to note is the following: In Figures 3.2*a* and 3.3*b*, the curves lie above the tangent to any point on them (they are said to *curve up* because they bend upward from the tangent), and the corresponding x components of acceleration are positive. The curves in Figures 3.2*b* and 3.3*a* lie below the tangent to any point; they are said to *curve down*, and the corresponding x components of acceleration are negative. So the sign of the x component of the acceleration is directly related to the curvature of an $x(t)$ curve:

The curvature of an $x(t)$ curve is a measure of the x component of acceleration. An upward curvature corresponds to a positive x component of acceleration. A downward curvature corresponds to a negative x component of acceleration.

3.4 (*a*) A car is speeding up in the negative x direction. In what direction do $\vec{a}$ and $\vec{v}$ point? (*b*) To which of the four graphs in Figures 3.2 and 3.3 does the situation represented in Figure 3.1*c* correspond?

3.2 Acceleration due to gravity

As an example of accelerated motion, think of an object falling in a straight line toward Earth's surface. We attribute the falling to *gravity* and say that there is a gravitational attraction between Earth and the object. Even though we don't know the ultimate cause of gravitational attraction, we can describe its effects.

CONCEPTS

Figure 3.4 Multiple-flash picture of a ball falling. The time interval between successive flashes is 0.05 s. Note how the distance traveled increases from one time interval to the next.

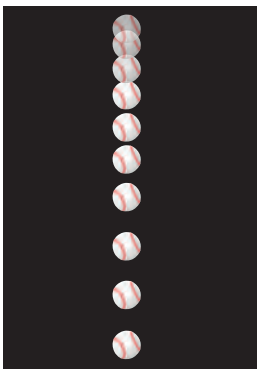


Figure 3.4 shows the positions of a falling ball recorded at equal time intervals of 0.05 s. Note how the spacing between successive positions increases as the ball drops. The picture tells us that the ball covers increasingly large distances in successive equal time intervals, and so we know that its speed increases as it falls. In other words, the ball accelerates. Two questions come to mind: *What is the magnitude of the acceleration as an object falls*? and *Does the acceleration depend on the physical characteristics of the object*?

Let us begin with the second question. If you drop a feather and a stone from the same height, there is no denying that the stone hits the ground first. This observation suggests that heavier objects fall faster than lighter ones—a common thought until the Italian mathematician, astronomer, and physicist Galileo Galilei (1564–1642) began to question its validity. Galileo observed that there is no difference between the amount of time taken by two stones to fall the same distance, even if the difference in weight is substantial and the stones are dropped from a considerable height. He argued that any observed difference in the amount of time it takes two objects to fall is a consequence of air resistance, which has a tendency to slow down lighter objects more than heavier ones.

3.5 Hold a book and a sheet of paper, cut to the same size as the book, side by side 1 m above the floor. Hold the paper parallel to the floor and the book with its covers parallel to the floor, and release them at the same instant. Which hits the floor first? Now put the paper on top of the book and repeat the experiment. What do you notice?

Figure 3.5 shows the effect of air resistance on a falling stone and a falling feather. One stone and one feather are dropped inside a glass tube from which the air has been pumped out. An identical stone and feather are released in air next to the tube. Figure 3.5*a* shows the situation at the simultaneous release of both pairs;

Figure 3.5 (*a*) Feathers and stones released from a certain height inside and outside an evacuated tube. (*b*) In air, the feather soon lags behind the stone; in vacuum, the two fall at the same speed.

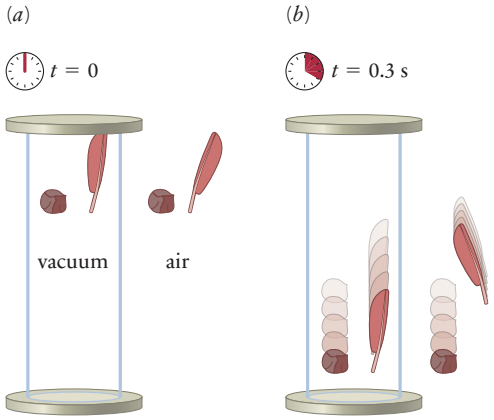


Figure 3.5*b* shows the situation 0.3 s after release. In air, the feather lags behind the stone. Inside the tube, however, both the stone and the feather fall the same distance. In vacuum, in other words, the feather "falls like a rock." Repeating this experiment with many different objects, we discover that, in the absence of air resistance, the acceleration due to gravity does not depend on the physical characteristics of the falling object.

The motion of objects moving under the influence of gravity only is called **free fall.** The feather inside the evacuated tube is in free fall, but the feather outside the tube is not because its motion is influenced not just by gravity but by air resistance, too.

Notice that the effect of air resistance on the stone outside the tube is hardly noticeable—it falls just as fast as the stone in vacuum. This is true for many objects as long as they are not too light and as long as they are not dropped from too great a height. So, unless you drop a light object from any height or a heavy one from the top of the Eiffel Tower, you can ignore air resistance and consider all dropped objects to be in free fall.

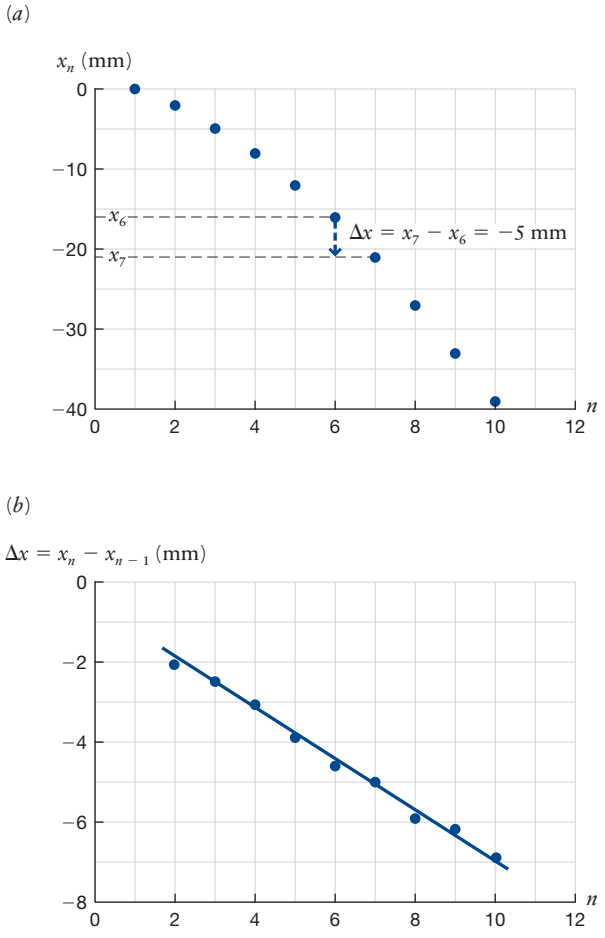
Let us now turn to the first question we posed earlier: What is the magnitude of the acceleration due to gravity? In **Figure 3.6*a*** I have plotted the positions of the ball (technically, its image) in Figure 3.4; for each image of the ball, I measured the position from the center of the ball to the center of the topmost ball. Figure 3.6*b* shows the x component of the displacement between successive exposures; note how the magnitude of the displacement increases from one exposure to the next, telling us that the ball is accelerating.

3.6 What does Figure 3.6*b* tell you about the velocity of the ball from one exposure to the next?

Checkpoint 3.6 is tricky not because I wanted to mislead you but because it is important to see the difference

CONCEPTS

Figure 3.6 (*a*) Position-versus-time graph of the freely falling ball (image) in Figure 3.4 for successive flashes $n = 1, 2, \ldots, 10$. (*b*) The change in the ball's x coordinate during the 10 exposures.

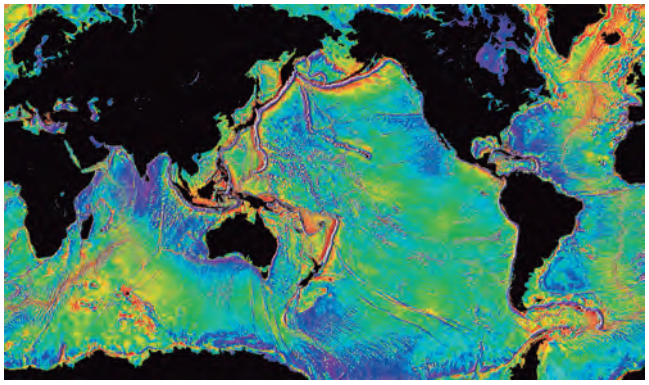


between plotting position versus time and plotting displacement (*change* in position) versus time.*

The ball's position at a given instant is simply its distance from the starting position, which is what Figure 3.6*a* shows. What Figure 3.6*b* shows is the x component of the *displacement* $\Delta x = x_f - x_i$ from one exposure to the next. Because the motion is not at constant velocity, the displacement changes from one exposure to the next. I include this graph to show the following remarkable fact: The data points lie on a straight line, which means that the magnitude of the displacement increases at a constant rate. Because velocity is proportional to displacement, this graph tells us that the velocity of the ball increases at a constant rate; in other words, the ball's acceleration is constant. So, for an object in free fall, the acceleration remains *constant* throughout the object's fall toward Earth.

*If you skipped the checkpoint, please do it now before I give you the answers. When you read a thriller, do you skip to the last page to find out how it ends? If you are puzzled by my questions, don't be discouraged and don't give up too soon. I wrote these questions precisely to arouse your curiosity. Discovering the answers to them is part of the learning process. Once you figure out the answers, you are unlikely to forget them. Taking the time to answer each checkpoint gives you something of lasting value. And if you like challenges, figuring out these checkpoints is without any doubt the easiest and most enjoyable way to learn physics. So take my advice to heart: Don't spoil the puzzle. You can do it—and you won't regret it!

Figure 3.7 Gravity map of Earth's oceans compiled from Seasat satellite data. The coloration shows variations in the value of the acceleration due to gravity of a few tenths of a percent.



Experiments show that, near Earth's surface in the absence of air resistance, the magnitude of the acceleration of all objects due to gravity for all objects in free fall is 9.8 m/s^2.

In the absence of air resistance, the magnitude of the acceleration of all falling objects is 9.8 m/s^2, and consequently the amount of time it takes to fall from a certain height is the same for all falling objects.

So, each second an object falls, its speed increases by about 10 m/s. After 1 s, an object released from rest near Earth's surface has a speed of about 10 m/s. After 2 s, its speed is about 20 m/s, and so on. The value of the acceleration due to gravity varies a few tenths of a percent over Earth's surface (**Figure 3.7**).

CONCEPTS

3.7 (*a*) Does the speed of a falling object increase or decrease? (*b*) If the positive x axis points up, does v_x increase or decrease as the object falls? (*c*) Is the x component of the acceleration positive or negative?

3.3 Projectile motion

So far we have considered only the motion of freely falling objects released from rest. Let us now see what happens when we *launch* an object upward, so that it is already moving when we release it. An object that is launched but not self-propelled is called a *projectile,* and its motion is called **projectile motion.** The path the object follows is called its **trajectory.** The motion of an object in free fall is a special case of projectile motion.

Consider a ball launched straight up, as in **Figure 3.8*a***. The ball first moves up and then reverses its direction of travel; its trajectory therefore is a straight line. Figure 3.8*b* shows a film or video clip of the motion of the ball; between frames 1 and 9 the ball moves vertically upward, and between frames 9 and 16 it moves vertically downward.

Figure 3.8 A ball launched vertically upward travels to its highest position and then retraces its path; the vertical "ruler" in each frame allows us to determine the position of the ball as a function of time.

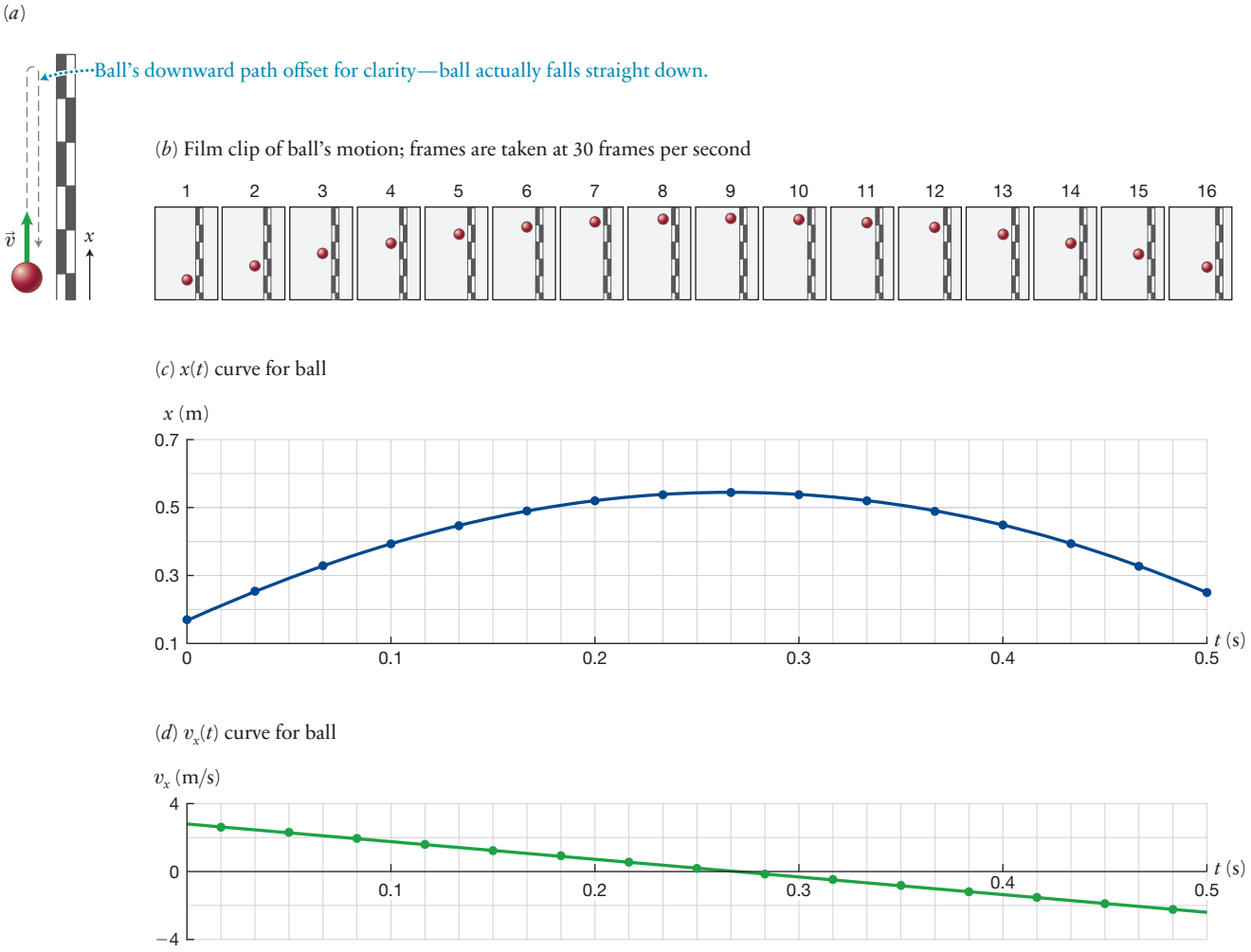


Figure 3.8*c* shows the ball's position for each frame in the clip, and in Figure 3.8*d* I have plotted the ball's average velocity in each frame, obtained by dividing the displacement from one frame to the next by the 33-ms time interval between frames. As you can see, the velocity gets smaller and smaller as the ball nears the top of the trajectory. The ball then reverses direction, and so its velocity becomes negative. Beyond frame 9, the ball's speed—the magnitude of its velocity—increases again.

Because the ball slows down as it moves upward, its acceleration $\vec{a}$ and velocity $\vec{v}$ must point in opposite directions. Thus, because the velocity points upward, the acceleration points downward while the ball is moving upward. As the ball moves downward, it speeds up, and so now its acceleration and velocity both point downward. In other words, the ball's acceleration points downward both on the way up and on the way down.

Figure 3.8*d* shows that the ball's velocity changes *linearly*—it changes by equal amounts in equal time intervals—and so the acceleration is constant. The slope of the straight line in Figure 3.8*d* is about $-10\ \text{m/s}^2$, which is the acceleration due to gravity. Therefore we see that, both on the way up and on the way down, the acceleration of a projectile is equal *both in magnitude and in direction* to that of a freely falling object.

If you throw a stone vertically upward and it has a velocity of +20 m/s right after leaving your hand, it slows down to +10 m/s after 1 s and reverses direction after 2 s. After 3 s, the stone is moving downward with a velocity of −10 m/s. Each second, gravity reduces the stone's velocity by another 10 m/s regardless of whether the stone is moving upward, moving downward, or instantaneously has zero velocity at the top. If you throw the stone harder, it leaves your hand with a higher speed, but once it leaves your hand, its motion is determined *solely* by gravity! In other words:

If an object is launched upward, the launch affects only the initial velocity. Once the object is released, the rest of its motion is determined by gravity alone (free fall).

3.8 Imagine throwing a ball downward so that it has an initial speed of 10 m/s. What is its speed 1 s after you release it? 2 s after?

CONCEPTS

What happens at the very top of the trajectory of a ball launched upward? Because the x component of the ball's velocity goes from positive to negative at the top, there must be an instant at which the x component of the velocity is zero. The x component of the acceleration, however, is *not* zero at the top. To see why, consider the motion over a small time interval Δt during which the ball reaches its highest position. Just before the ball reaches the top, its velocity $\vec{v}$ points upward; just after the ball reaches the top, the velocity points downward. The vector $\Delta\vec{v}$ therefore points downward, and so does the acceleration $\vec{a}$. This holds true regardless of the size of the time interval: As we make Δt smaller, Δv_x becomes smaller, too, but the ratio $\Delta v_x/\Delta t$ remains equal to the slope of the $v_x(t)$ curve in Figure 3.8*d*. So the acceleration of the ball is the same *everywhere*: on the way up, on the way down, and at the top.

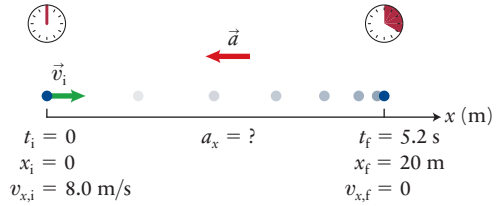
If it seems odd that the acceleration at the top is nonzero, remember that the velocity must change there from up to down, and any change in velocity—be it a change in magnitude or a change in direction—means a nonzero acceleration. That the x component of the velocity is zero doesn't matter: The x component of the velocity doesn't remain zero for any finite time interval; it is only *instantaneously* zero, and in order for it to change, there must be an acceleration. Note, also, that even though the x component of the velocity is instantaneously zero, it is not correct to say that the ball is "instantaneously at rest"; *at rest* means that the position of the ball is not changing (see Section 2.1), which is not the case at any position along the ball's trajectory, including the very top.

3.4 Motion diagrams

To help us visualize the motion of an object described in a problem, we make a **motion diagram** representing the position of the moving object at equally spaced time intervals. Motion diagrams look very similar to the multiple-flash picture shown in Figure 3.4 and provide a visual overview of all you know about the motion. The procedure for making such diagrams is explained in the box on page 93.

Suppose a bicycle rider going 8.0 m/s applies the brake and slows down to a stop over a distance of 20 m in 5.2 s. The motion diagram in **Figure 3.9** shows that same information in graphical form. Note how the diagram shows all we know about the motion: the initial velocity, the distance it takes to come to a stop, and the time interval over which the stopping takes place. The spacing of the dots shows that the object of interest is slowing down. The question mark shows what information is missing. Making such a diagram helps you visualize the problem, organize the information given, and choose a convenient reference axis.

Figure 3.9 Motion diagram for bicycle rider slowing down to a stop.



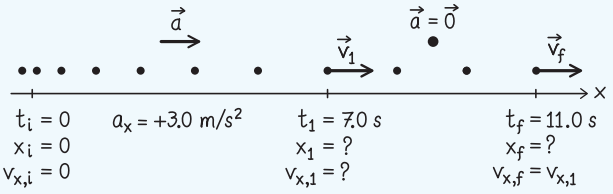
Nearly all problems in physics—and motion diagrams in particular—involve considering changes that occur over a certain time interval. Throughout this book we shall use clocks as in Figure 3.9 to mark the beginning and end of the time interval of interest. An important part of any physics problem is to record the parameters that characterize the situation under consideration at the beginning and end of that time interval (the *initial* and *final conditions*). When drawing motion diagrams, for example, we record an object's initial and final positions as well as its initial and final velocities. We begin applying this procedure in the next two examples.

EXERCISE 3.1 Accelerating car

Starting from rest, a car accelerates for 7.0 s at 3.0 m/s^2 and then travels at constant velocity for another 4.0 s. Make a motion diagram representing this motion.

SOLUTION The motion consists of two parts: constant acceleration during the first 7.0 s and constant velocity (zero acceleration) during the last 4.0 s. In my motion diagram (**Figure 3.10**) I show both parts of the motion: On the left the spacing between the dots increases to show the acceleration; on the right the spacing is constant to show motion at constant velocity.

Figure 3.10



I chose a positive x axis pointing in the direction of travel and with the origin at the car's starting position. The initial conditions are $t_i = 0$, $x_i = 0$, and $v_{x,i} = 0$.

Before specifying the final conditions, I must consider the instant where the car stops accelerating. This instant, which I denote by t_1, occurs 7.0 s after the car begins to accelerate. Neither the car's position nor its velocity at that instant is known, so I denote these values by question marks in the diagram.

The car then moves at constant velocity for 4.0 s, so the final conditions are $t_f = 11.0$ s, $x_f = ?$, and $v_{x,f} = v_{x,1}$. I complete the diagram by indicating that the x component of the car's acceleration is +3.0 m/s^2 between t_i and t_1 and zero between t_1 and t_f. ✔

✔ The checkmark shows that I carried out the following checks: Vectors and scalars used appropriately; Everything answered; No unknowns left; Units correct; Significant digits justified.

CONCEPTS

PROCEDURE: Analyzing motion using motion diagrams

When you solve motion problems, it is important to begin by making a diagram that summarizes what you know about the motion.

1. Use dots to represent the moving object at equally spaced time intervals. If the object moves at constant speed, the dots are evenly spaced; if the object speeds up, the spacing between the dots increases; if the object slows down, the spacing decreases.
2. Choose an x (position) axis that is convenient for the problem. Most often this is an axis that (*a*) has its origin at the initial or final position of the object and (*b*) is oriented in the direction of motion or acceleration.
3. Specify the position and velocity at all relevant instants. In particular, specify the *initial conditions*—position and velocity at the beginning of the time interval of interest—and the *final conditions*—position and velocity at the end of that time interval. Also specify all positions where the velocity reverses direction or the acceleration changes. Label any unknown parameters with a question mark.
4. Indicate the acceleration of the object between all the instants specified in step 3.
5. To consider the motion of more than one object, draw separate diagrams side by side, one for each object, using one common x axis.
6. If the object reverses direction, separate the motion diagram into two parts, one for each direction of travel.

EXAMPLE 3.2 Can this be?

A newspaper article you read claims that by the time it reaches the ground, a stone dropped from the top of the Empire State Building (which has approximately 100 floors) "travels the length of a window faster than you can say *Watch out! A stone!*" Estimate whether this is true.

1 GETTING STARTED At first this problem appears ill-defined. What am I supposed to calculate? Where to begin? If the answer to the first question is not immediately clear, the answer to the second is: with a motion diagram. I therefore begin by organizing the information following the procedure in the Procedure box "Analyzing motion using motion diagrams." I show the motion represented by a series of dots of increasing spacing to represent the acceleration due to gravity (**Figure 3.11**). I chose an x axis having its origin at the top of the building (the stone's initial position) and pointing downward (in the direction of the motion). This choice of axis means that downward corresponds to the *positive* x direction—x increases as the stone falls. The initial conditions are then

$$t_i = 0, \quad x_i = 0, \quad v_{x,i} = 0.$$

The clock reading t_i and position x_i are zero because of my arbitrary choice of reference point. I could have chosen any other values, but the choice I made simplifies things later on. That the x component of the initial velocity is zero is an assumption because the problem doesn't specify this detail. All I know is that the stone is dropped, and it seems reasonable to assume it is not thrown downward or upward, but simply released from rest.

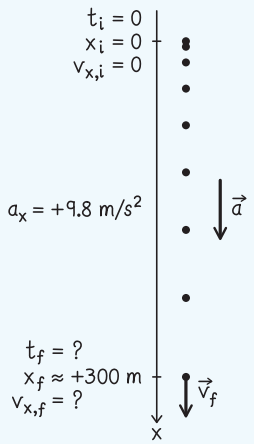
Because I am interested in the stone's speed just as it reaches the ground, I choose that instant as the final instant. To estimate the distance between the initial and final positions, I assume that each floor is about 3 m high, making the building about 300 m high.* So the final conditions are

$$t_f = ?, \quad x_f = +300 \text{ m}, \quad v_{x,f} = ?.$$

Between $x_i = 0$ and $x_f = +300$ m, the stone accelerates downward because of gravity. I make a simplifying assumption that air resistance doesn't affect the acceleration of the stone very much, so I write "$a_x = +9.8 \text{ m/s}^2$" between the initial and final positions.†

(Continued)

Figure 3.11



*The Empire State Building is 381 m high, and so my estimate is off by about 30%, but you will soon see that this error is irrelevant.

†Note that a_x is positive because both the acceleration and the x axis point downward. This assumption is almost certainly an oversimplification because air resistance cannot be ignored for a fall from about 300 m. However, we have not yet developed the tools to include air resistance, and so I have no other choice.

CONCEPTS

2 DEVISE PLAN To complete the estimating task given in the problem statement, I need to compare the speed of the stone at the bottom of its fall to the speed it would have to have if it traveled the length of a window in the same time interval it takes to say "Watch out! A stone!" Because no values are given, I'll have to use estimates and then evaluate the speed using the relation among speed, distance traveled, and acceleration.

3 EXECUTE PLAN I estimate that it takes roughly 2 s to say those words and a window is about 2 m tall. So, as long as the stone is traveling faster than $(2\text{ m})/(2\text{ s}) = 1\text{ m/s}$ just before it hits the ground, it passes by a window before the four words can be uttered. I also know that the acceleration in free fall is about 10 m/s^2, so the speed of the stone increases by 10 m/s each second. That means that the stone's speed is 10 m/s 1 s after it is released. I know from experience that it takes far longer than 1 s for an object to fall such a great distance, and so, even without further calculation, I see that the statement must be true. ✔

4 EVALUATE RESULT Even if I account for some slowing down due to air resistance, I know from experience that a stone dropped from even a much smaller height travels more than 1 m in 1 s just before it reaches the ground. (In the second part of this chapter, you will be able to use the motion diagram to calculate how fast the stone moves just before hitting the ground.)

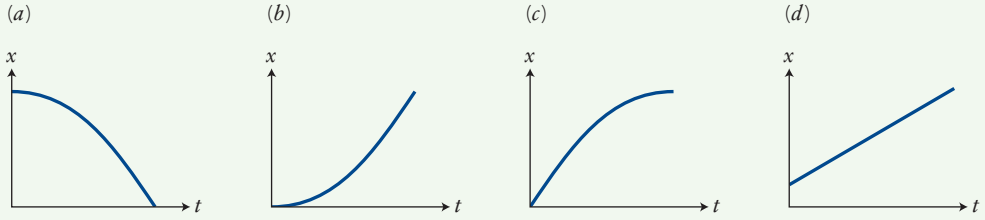
3.9 Make a motion diagram for the following situation: A seaside cliff rises 30 m above the ocean surface, and a person standing at the edge of the cliff launches a rock vertically upward at a speed of 15 m/s. After reaching the top of its trajectory, the rock falls into the water.

CONCEPTS

Self-quiz

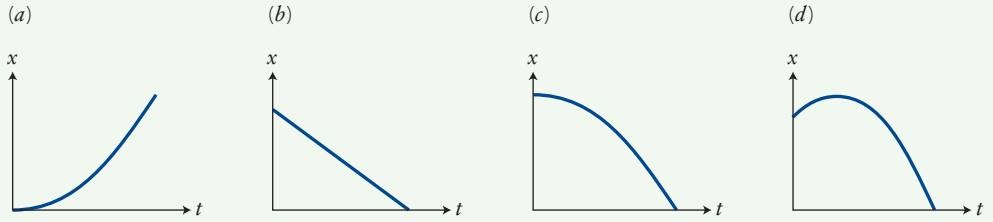
1. Two stones are released from rest at a certain height, one 1 s after the other. (*a*) Once the second stone is released, does the difference in their speeds increase, decrease, or stay the same? (*b*) Does their separation increase, decrease, or stay the same? (*c*) Is the time interval between the instants at which they hit the ground less than, equal to, or greater than 1 s? (Use $x(t)$ curves to help you visualize this problem.)
2. Which of the graphs in **Figure 3.12** depict(s) an object that starts from rest at the origin and then speeds up in the positive x direction?

Figure 3.12



3. Which of the graphs in **Figure 3.13** depict(s) an object that starts from a positive position with a positive x component of velocity and accelerates in the negative x direction?

Figure 3.13



4. Draw a motion diagram for an object that initially has a negative x component of velocity but has a positive x component of acceleration. (Include motion in one direction only.)

Answers

1. (*a*) Both stones accelerate at about 10 m/s^2, and so the speed of each increases about 10 m/s each second. Because the speeds increase at the same rate, the difference in the speeds remains the same. (*b*) As indicated by the vertical lines between the $x(t)$ curves for both stones in **Figure 3.14**, the separation increases because the speed of the first stone is always greater. (*c*) Both stones execute the same motion—they start from rest and have the same acceleration and travel the same distance. So the second stone always remains 1 s behind (the $x(t)$ curves are separated horizontally by 1 s at all instants; see the horizontal lines separating the curves in Figure 3.14). Therefore the second stone hits the ground 1 s after the first one hits.
2. Graph *b* is the correct answer because its initial position is zero and the slope is initially zero but then increasing, indicating that the object speeds up. Graph *a* depicts a motion that does not begin at the origin and has the object speeding up in the negative x direction. Graph *c* depicts an object slowing down while moving in the positive x direction. Graph *d* depicts an object that is initially not at the origin and that moves at constant velocity in the positive x direction.
3. Graph *d*. Graph *a* does not have a positive initial position. Graph *b* represents zero acceleration. Graph *c* represents zero initial velocity.
4. An object that has a negative x component of velocity must slow down in order to have a positive x component of acceleration. **Figure 3.15** shows what a motion diagram for this situation looks like.

Figure 3.14

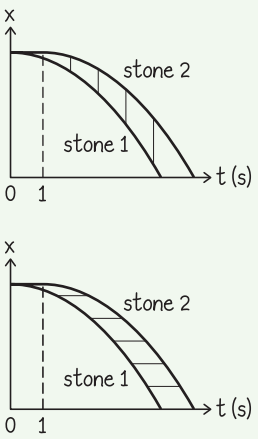
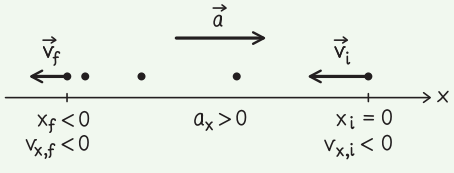


Figure 3.15



3.5 Motion with constant acceleration

As we saw in Section 3.1, the x component of the average acceleration is the change in the x component of the velocity divided by the time interval during which this change took place:

$$a_{x,\text{av}} \equiv \frac{\Delta v_x}{\Delta t} = \frac{v_{x,\text{f}} - v_{x,\text{i}}}{t_\text{f} - t_\text{i}}. \tag{3.1}$$

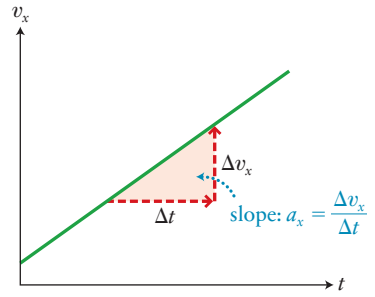
Notice the similarity between this definition and the definition of average velocity (Eq. 2.14):

$$v_{x,\text{av}} \equiv \frac{\Delta x}{\Delta t} = \frac{x_\text{f} - x_\text{i}}{t_\text{f} - t_\text{i}}. \tag{3.2}$$

Velocity is the rate of change in position, and acceleration is the rate of change in velocity.

Let us begin by analyzing motion with constant acceleration. The x component of the velocity of an object whose acceleration is constant increases by equal amounts Δv_x during equal time intervals Δt. For example, the x component of the velocity of an object that has a constant acceleration $a_x = +5\ \text{m/s}^2$ increases by 5 m/s each second. If we plot the object's x component of velocity versus time, we therefore get a straight line with a nonzero slope, as in **Figure 3.16**.

Figure 3.16 For an object moving with constant acceleration, the $v_x(t)$ curve is a straight line.



To obtain an expression that allows us to determine the x component of the final velocity $v_{x,\text{f}}$ at some final instant t_f, we rewrite Eq. 3.1 in the form*

$$\Delta v_x \equiv v_{x,\text{f}} - v_{x,\text{i}} = a_x \Delta t \quad \text{(constant acceleration)} \tag{3.3}$$

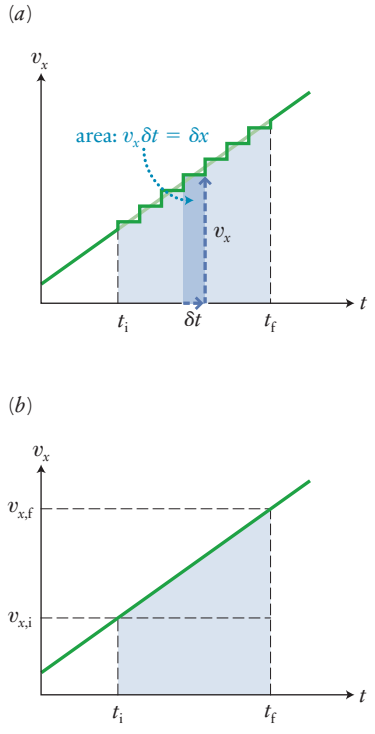
or

$$v_{x,\text{f}} = v_{x,\text{i}} + a_x \Delta t \quad \text{(constant acceleration)}. \tag{3.4}$$

So, if we know the x component of the initial velocity $v_{x,\text{i}}$ at some initial instant t_i for an object moving with constant acceleration, we add $a_x \Delta t$ to this x component of the velocity to obtain the x component of the final velocity $v_{x,\text{f}}$ at a later instant $t_\text{f} = t_\text{i} + \Delta t$.

How do we determine the displacement Δx for an accelerating object? In Section 2.8 we learned that the displacement of an object moving at constant velocity is equal to the area under the $v_x(t)$ curve. The motion we are considering now is not at constant velocity because acceleration is involved, but we can still apply what we know. We begin by approximating the $v_x(t)$ curve of Figure 3.16 by the stepped curve shown at the top of the shaded area in **Figure 3.17a**. The shaded area consists of many small rectangles of equal width δt adjusted in height to fit under the stepped $v_x(t)$ curve. (We use a lowercase Greek delta, δ, instead of an uppercase delta, Δ, to indicate that the time interval is small.) The area of each rectangle gives the displacement δx during an interval δt, and this area is the x component of the velocity at the beginning of the time interval multiplied by δt. The total displacement $\Delta x = x_\text{f} - x_\text{i}$ in the time interval from t_i to t_f is approximately equal to the sum of the areas of all the rectangles. The smaller the width of the rectangles, the more accurate the approximation.

Figure 3.17 Determining the displacement for an object with constant acceleration.



*Because the acceleration is constant, it always has the same (average) value, and so we can drop the subscript av.

QUANTITATIVE TOOLS

In the limit that Δt approaches zero, the displacement $\Delta x = x_f - x_i$ is given exactly by the trapezoidal area under the $v_x(t)$ curve between t_i and t_f in Figure 3.17*b*. This area is equal to the sum of the areas of the shaded triangle and rectangle in Figure 3.17*b*. The area of the triangle is

$$\tfrac{1}{2}(v_{x,f} - v_{x,i})(t_f - t_i) = \tfrac{1}{2}\Delta v_x\, \Delta t = \tfrac{1}{2}\, a_x(\Delta t)^2, \tag{3.5}$$

where I have used Eq. 3.3 to substitute $a_x \Delta t$ for Δv_x. The area of the rectangle is

$$(v_{x,i} - 0)(t_f - t_i) = v_{x,i}\Delta t. \tag{3.6}$$

The displacement is thus the sum of Eqs. 3.5 and 3.6:

$$x_f - x_i = v_{x,i}\Delta t + \tfrac{1}{2}a_x(\Delta t)^2 \quad \text{(constant acceleration)}, \tag{3.7}$$

and the object's final position is

$$x_f = x_i + v_{x,i}\Delta t + \tfrac{1}{2}a_x(\Delta t)^2 \quad \text{(constant acceleration)}. \tag{3.8}$$

Equations 3.4 and 3.8 can be written without the delta notation when we reset our clock to zero at the beginning of a motion segment, so that $t_i = 0$ and $\Delta t = t_f - t_i = t_f$. Equations 3.8 and 3.4 then become, respectively,

$$x_f = x_i + v_{x,i}t_f + \tfrac{1}{2}a_x t_f^2 \quad \text{(constant acceleration)} \tag{3.9}$$

and

$$v_{x,f} = v_{x,i} + a_x t_f \quad \text{(constant acceleration)}, \tag{3.10}$$

giving us the position and the x component of the velocity of an object at any instant t_f.

Equations 3.9 and 3.10 are the basic equations for motion at constant acceleration when the initial conditions at $t_i = 0$ are known. Note that when $a_x = 0$, these equations become identical to the equations for motion at constant velocity: Eq. 3.10 yields an x component of the velocity v_x that is constant, and Eq. 3.9 becomes identical to Eq. 2.19.

Equation 3.9 gives the final position of an object at instant t_f. Because this instant is arbitrary, we can drop the subscript f to obtain an expression for the $x(t)$ curve of an object moving at constant acceleration:

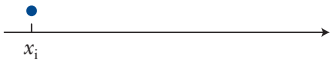
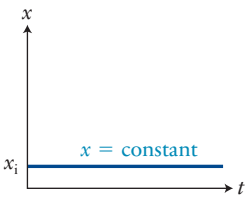
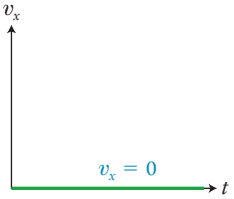
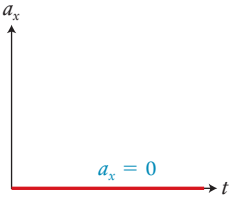
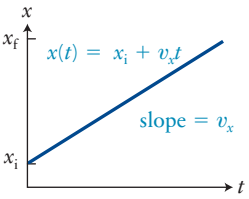
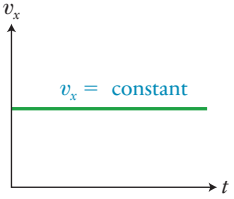
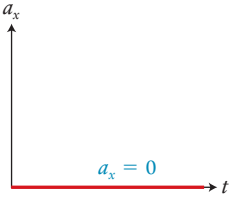
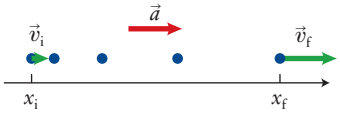
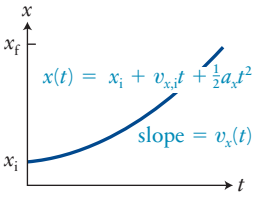
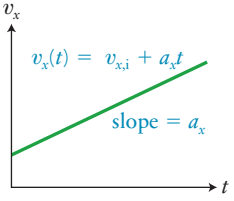
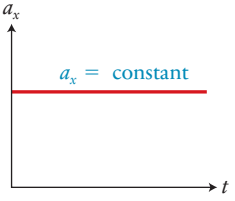
$$x(t) = x_i + v_{x,i}t + \tfrac{1}{2}a_x t^2 \quad \text{(constant acceleration)}. \tag{3.11}$$

This expression shows that, for motion with constant acceleration, position depends on the square of t, and the $x(t)$ curve is thus a parabola. If a_x is positive, the parabola curves upward; if it is negative, the parabola curves downward (for an example, see Figure 3.8*c*). Likewise, from Eq. 3.10 we see that the $v_x(t)$ curve for motion at constant acceleration is a straight line:

$$v_x(t) = v_{x,i} + a_x t \quad \text{(constant acceleration)}. \tag{3.12}$$

For motion at constant velocity, we can set $a_x = 0$ in Eqs. 3.11 and 3.12, yielding straight lines for both $x(t)$ and $v_x(t)$. **Table 3.1** summarizes the three basic types of motions we have studied so far: at rest, constant velocity, and constant acceleration.

Table 3.1 Kinematics graphs for three basic types of motion

Motion diagram	Position versus time	Velocity versus time	Acceleration versus time
At rest 			
Constant velocity 			
Constant acceleration 			

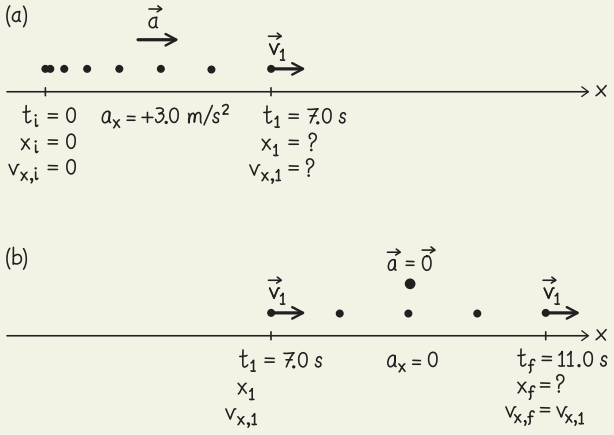
EXAMPLE 3.3 Accelerating car, again

Use Eqs. 3.11 and 3.12 to determine how far the car in Exercise 3.1 travels.

❶ **GETTING STARTED** To quantitatively analyze the situation in Exercise 3.1, I begin by separating the motion diagram in Figure 3.10 into two parts: one representing the acceleration of the car (**Figure 3.18*a***) and one representing its motion at constant speed (Figure 3.18*b*). Note that the initial position and velocity in Figure 3.18*b* are not labeled with a question mark; even though they are not given, they are identical to the final position and velocity in Figure 3.18*a*, and so their values can be determined by solving the first part of the problem. The only unknown in the second part is the final position of the car.

❷ **DEVISE PLAN** Equation 3.11 allows me to determine the position of the accelerating car during the first part of the motion. To calculate the distance traveled during the second part of the motion, I need to know the velocity of the car. I can obtain this velocity from Eq. 3.12. From Eq. 2.18, I know that the distance traveled is then the product of this constant velocity and the duration of the time interval (4.0 s). Adding this distance to the position of the car at the end of the first part of the motion then gives me the total distance traveled by the car.

Figure 3.18



❸ **EXECUTE PLAN** Given that $x_i = 0$ and $v_{x,i} = 0$, Eq. 3.11 reduces to $x_1 = x(t_1) = \frac{1}{2}a_x t_1^2$. Substituting $a_x = +3.0\ \text{m/s}^2$ and $t_1 = 7.0\ \text{s}$, I obtain for the position of the car at the end of the first part of the motion:

$$x_1 = \tfrac{1}{2}(+3.0\ \text{m/s}^2)(7.0\ \text{s})^2 = +74\ \text{m}.$$

QUANTITATIVE TOOLS

From Eq. 3.12, I obtain the velocity at the beginning of the second interval:

$$v_1 = 0 + (+3.0\ \text{m/s}^2)(7.0\ \text{s}) = +21\ \text{m/s}.$$

At this constant velocity, the car's displacement during the 4.0 s of the second interval is

$$\Delta x = v_{x,1}\Delta t = (+21\ \text{m/s})(11.0\ \text{s} - 7.0\ \text{s}) = +84\ \text{m},$$

and so the x component of the car's total displacement is $(+74\ \text{m}) + (+84\ \text{m}) = +158\ \text{m}$. The car in Exercise 3.1 thus traveled a distance of 158 m. ✔

4 EVALUATE RESULT With an acceleration of $3.0\ \text{m/s}^2$ (about one-third of the acceleration due to gravity), the car reaches a speed of

$$\frac{21\ \text{m}}{1\ \text{s}} \times \frac{1\ \text{mi}}{1609\ \text{m}} \times \frac{3600\ \text{s}}{1\ \text{h}} = 47\ \text{mi/h}$$

over the 7.0 s of the first interval and travels a distance of 74 m or about the length of a sports field. Both the car's final speed and the distance traveled are in line with what I would expect. During the second part of the motion, the car travels roughly the same distance in 4.0 s, which is also reasonable.

EXAMPLE 3.4 Collision or not?

You are bicycling at a steady 6.0 m/s when someone suddenly walks into your path 2.5 m ahead. You immediately apply the brakes, which slow you down at $6.0\ \text{m/s}^2$. Do you stop in time to avoid a collision?

1 GETTING STARTED In order to avoid a collision, you must come to a stop in less than 2.5 m. The problem therefore boils down to calculating the distance traveled under the given conditions and comparing this distance to 2.5 m.

2 DEVISE PLAN Equation 3.7 gives me the displacement, but I don't know the time interval Δt. I can obtain this time interval from Eq. 3.4: $\Delta t = (v_{x,\text{f}} - v_{x,\text{i}})/a_x$, which contains no unknowns on the right side.

3 EXECUTE PLAN Substituting the expression for the time interval from Eq. 3.4 into Eq. 3.7 gives the x component of the displacement necessary to stop:

$$\Delta x = v_{x,\text{i}}\frac{v_{x,\text{f}} - v_{x,\text{i}}}{a_x} + \tfrac{1}{2}a_x\left(\frac{v_{x,\text{f}} - v_{x,\text{i}}}{a_x}\right)^2 = \frac{v_{x,\text{f}}^2 - v_{x,\text{i}}^2}{2a_x}. \qquad (1)$$

When I let the x axis point in the direction of the motion, the x component of the initial velocity is $v_{x,\text{i}} = +6.0\ \text{m/s}$, the x component of the final velocity is zero, $v_{x,\text{f}} = 0$, and the x component of the acceleration is $a_x = -6.0\ \text{m/s}^2$. Substituting these values into Eq. 1, I get

$$\Delta x = \frac{0 - (+6.0\ \text{m/s})^2}{2(-6.0\ \text{m/s}^2)} = +3.0\ \text{m},$$

which is more than the 2.5 m separating you from the person walking into your path. You will not be able to stop in time to avoid a collision. ✔

4 EVALUATE RESULT Because $a_x = -6.0\ \text{m/s}^2$, your speed decreases by 6.0 m/s each 1 s, and so it takes 1 s to come to a stop from your initial speed of 6.0 m/s. Because your speed is decreasing, you travel less than 6.0 m, but without carrying out the calculation, I can't tell whether or not that distance is less than 2.5 m. All I can really say is that my answer should be less than 6.0 m, which it is.

Equation 3.4 allows us to determine the x component of the final velocity of an object undergoing constant acceleration after a certain time interval Δt. By multiplying both sides of Eq. 1 in Example 3.4 by $2a_x$ and rearranging terms, I get an expression for the x component of the final velocity of an object undergoing constant acceleration over a certain displacement Δx:

$$v_{x,\text{f}}^2 = v_{x,\text{i}}^2 + 2a_x\Delta x \quad \text{(constant acceleration)}. \qquad (3.13)$$



3.10 Determine the velocity of the stone dropped from the top of the Empire State Building in Example 3.2 just before the stone hits the ground.

3.6 Free-fall equations

The magnitude of the **acceleration due to gravity** is designated by the letter g:

$$g \equiv |\vec{a}_{\text{free fall}}|. \qquad (3.14)$$

As noted in Section 3.2, near Earth's surface the magnitude of the acceleration due to gravity is $g = 9.8\ \text{m/s}^2$. The direction of this acceleration is

always downward, and so if we choose a positive x axis that points upward, the x component of the acceleration due to gravity is given by $a_x = -g$, where the minus sign takes into account that the acceleration vector points downward.

If an object is dropped from a certain height with zero initial velocity along an upward-pointing x axis, then, according to Eqs. 3.9 and 3.10,

$$x_f = x_i - \tfrac{1}{2}gt_f^2 \tag{3.15}$$

and

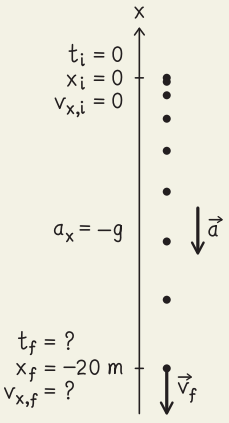
$$v_{x,f} = -gt_f. \tag{3.16}$$

EXAMPLE 3.5 Dropping the ball

Suppose a ball is dropped from height $h = 20$ m above the ground. How long does it take to hit the ground, and what is its velocity just before it hits?

❶ **GETTING STARTED** I begin by drawing a motion diagram, arbitrarily choosing an x axis that points upward and has its origin at the initial position of the ball (Figure 3.19). I make two simplifying assumptions: that the ball is released from rest ($v_{x,i} = 0$ at $t_i = 0$) and that I may ignore air resistance (so that the ball is in free fall).

Figure 3.19



When the ball is released from rest at the origin, the initial conditions are

$$t_i = 0, \quad x_i = 0, \quad v_{x,i} = 0.$$

The final position x_f at instant t_f is a distance h below the initial position, and so, just before impact at instant t_f, the final conditions are

$$t_f = ?, \quad x_f = -h, \quad v_{x,f} = ?.$$

As the ball falls, its acceleration is negative (because $\vec{a}$ points toward increasingly negative values of x), and so $a_x = -g$.

❷ **DEVISE PLAN** Because the acceleration of the ball is constant, I can use Eqs. 3.9 and 3.10. Doing so gives me two equations to determine the two unknowns t_f and $v_{x,f}$.

❸ **EXECUTE PLAN** Substituting the initial and final conditions into Eq. 3.9 yields

$$-h = 0 + 0 - \tfrac{1}{2}gt_f^2 = -\tfrac{1}{2}gt_f^2,$$

and so

$$t_f = \sqrt{\frac{2h}{g}}.$$

Substituting $h = 20$ m and $g = 9.8$ m/s^2, I obtain for the amount of time the ball takes to fall 20 m

$$\Delta t = t_f - t_i = \sqrt{\frac{2h}{g}} - 0 = \sqrt{\frac{2(20\text{ m})}{9.8\text{ m/s}^2}} = \sqrt{4.0\text{ s}^2}$$

$$= 2.0\text{ s.} \checkmark \tag{1}$$

Next I use Eq. 3.10. Because the ball starts from rest, I have

$$v_{x,f} = 0 - gt_f = -gt_f = -(9.8\text{ m/s}^2)(2.0\text{ s})$$

$$= -20\text{ m/s.} \checkmark$$

❹ **EVALUATE RESULT** The 2.0 s I obtained for the time interval it takes the ball to fall is not unreasonable: 20 m corresponds approximately to the height of a six-story building, and from experience I know that something dropped from such a height takes a few seconds to reach the ground. My answer for the final velocity $v_{x,f} = -20$ m/s also makes sense: The x component of the velocity is negative because it points in the negative x direction, and if the ball were moving always at a constant speed of 20 m/s, it would cover the 20-m distance in 1 s. Because it moves at that speed only at the end of the drop, it takes longer to fall, as Eq. 1 shows.

3.11 Repeat Example 3.5 for an x axis that points downward (leaving the origin at the ball's initial position).

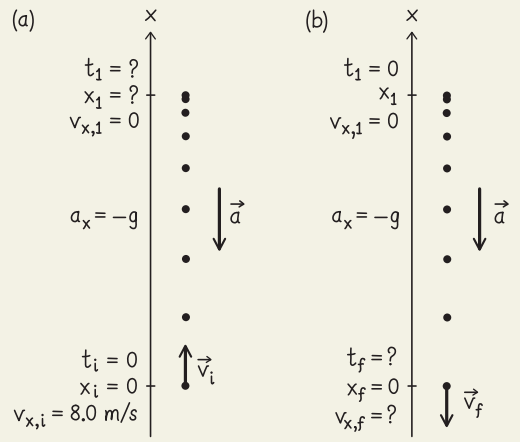
QUANTITATIVE TOOLS

EXAMPLE 3.6 A stone launched up

A stone is launched straight up from ground level at a speed of 8.0 m/s. (*a*) How high does it rise? (*b*) How many seconds does it take for the stone to hit the ground?

❶ **GETTING STARTED** I am asked to consider two parts of the trajectory: part *a* from the launch position to the top of the trajectory and part *b* from the launch position to the top and back down to the launch position. I begin therefore by drawing two motion diagrams: one for the upward motion and the other for the downward motion. I (arbitrarily) choose an x axis pointing upward, put the origin at the launching position, and start the clock at $t_i = 0$ at the instant the stone is launched (**Figure 3.20**).

Figure 3.20



For the upward motion (Figure 3.20*a*), the initial conditions are $t_i = 0$, $x_i = 0$, and $v_{x,i} = +8.0$ m/s. The hard part of this problem is specifying the final conditions, and here is where the motion diagram helps. I know neither the instant t_1 at which the highest position is reached nor the location x_1 of that position. I do, however, know *something* about the stone at the highest position. (Put the book down and think for a second—you know it, too.)

The key is the word *highest:* The stone reverses direction at the highest position, and so its velocity must be zero there. So, for the upward part of the motion, the stone's final conditions are $t_f = t_1 = ?$, $x_f = x_1 = ?$ and $v_{x,f} = v_{x,1} = 0$.

For the downward motion, I reset the clock to zero at the top, as indicated by $t_1 = 0$ in Figure 3.20*b*. Doing this simplifies the calculation because it makes this part of the problem nearly identical to Example 3.5. I should not forget, however, that $\Delta t = t_f - t_i = t_f - 0 = t_f$ now represents the duration of the downward motion only, not the total duration for the upward and downward travel.

❷ **DEVISE PLAN** To determine the height reached by the stone, I need to obtain a value for x_1. The stone is undergoing constant acceleration, so its position and velocity are given by Eqs. 3.8 and 3.4. For the upward part of the motion I see from my motion diagram that there are two unknowns (x_1 and t_1), and so the number of unknowns is equal to the number of equations. After I obtain x_1 and t_1, the downward part of the motion also contains two unknowns ($v_{x,f}$ and t_f), which I can calculate using the same two equations.

❸ **EXECUTE PLAN** (*a*) For the upward portion of the motion I can use Eqs. 3.9 and 3.10 instead of Eqs. 3.8 and 3.4 because the clock reads zero at the beginning of the motion. With the x axis pointing upward, the acceleration of the stone is $a_x = -g$, and so these equations become

$$x_1 = 0 + v_{x,i}t_f - \tfrac{1}{2}gt_f^2 \tag{1}$$

and

$$0 = v_{x,i} - gt_f. \tag{2}$$

Solving Eq. 2 for t_f, I get

$$t_f = \frac{v_{x,i}}{g} = \frac{+8.0\text{ m/s}}{9.8\text{ m/s}^2} = 0.82\text{ s}. \tag{3}$$

Substituting this result into Eq. 1, I determine the height to which the stone rises

$$x_1 = v_{x,i}\left(\frac{v_{x,i}}{g}\right) - \tfrac{1}{2}g\left(\frac{v_{x,i}}{g}\right)^2 = \frac{v_{x,i}^2}{g} - \frac{v_{x,i}^2}{2g} = \frac{v_{x,i}^2}{2g}, \tag{4}$$

and so

$$x_1 = \frac{(+8.0\text{ m/s})^2}{2(9.8\text{ m/s}^2)} = +3.3\text{ m}. \checkmark$$

(*b*) For the downward part of the motion, Eq. 3.7 becomes

$$0 - x_1 = 0 - \tfrac{1}{2}g(t_f - t_1)^2.$$

Rearranging terms and substituting Eq. 4 for x_1, I have

$$(t_f - t_1)^2 = \frac{2x_1}{g} = \frac{2v_{x,i}^2}{2g^2} = \left(\frac{v_{x,i}}{g}\right)^2,$$

which shows that the amount of time it takes for the stone to fall back to the ground, $\Delta t = t_f - t_1 = v_{x,i}/g$, is identical to the amount of time it takes for the stone to reach its highest position (see Eq. 3).

The amount of time it takes for the round trip is

$$t_f = 2\left(\frac{v_{x,i}}{g}\right) = 2\left(\frac{8.0\text{ m/s}}{9.8\text{ m/s}^2}\right) = 1.6\text{ s}. \checkmark$$

❹ **EVALUATE RESULT** That it takes as much time to travel up as it takes to fall back down makes sense if you look at the position-versus-time graph for a projectile in Figure 3.8*c*. The parabola is symmetrical about the maximum, meaning that the acceleration is the same in both directions, and so both parts of the motion must take the same amount of time. The stone's initial velocity is 8.0 m/s, so at an acceleration of about 10 m/s^2, it should take 0.8 s to slow down to zero and reach its highest position. The downward part of the motion is similar to the upward part, with the motion reversed, and so the total time should be twice as long. At a constant 8.0 m/s, the stone would travel (0.8 s)(8.0 m/s) = 6.4 m, but because it slows it travels only half that distance. My answers appear to make sense!

3.12 In Example 3.6, what is the x component of the velocity of the stone just before it hits the ground?

3.7 Inclined planes

Galileo concluded that the acceleration due to gravity is constant even though he was unable to study falling objects using high-speed video or multiple-flash photography. Instead, he used inclined planes at shallow angles of incline to reduce the acceleration. By rolling balls along such inclined planes, he came to the following remarkable conclusion:

When a ball rolls down an incline starting from rest, the ratio of the distance traveled to the square of the amount of time needed to travel that distance is constant.

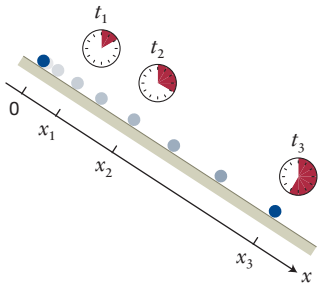


Figure 3.21 Galileo measured how long it takes a ball to roll down planes inclined at different angles and concluded that the acceleration due to gravity is constant.

For instance, if a ball released from position $x_i = 0$ at $t_i = 0$ reaches position x_1 at instant t_1, x_2 at t_2, and x_3 at t_3 (**Figure 3.21**), Galileo found experimentally that

$$\frac{x_1}{t_1^2} = \frac{x_2}{t_2^2} = \frac{x_3}{t_3^2}. \tag{3.17}$$

Note that for an object moving with constant acceleration and starting from rest at $x_i = 0$ and $t_i = 0$, we have, from Eq. 3.9,

$$x_f = \tfrac{1}{2} a_x t_f^2, \tag{3.18}$$

so that

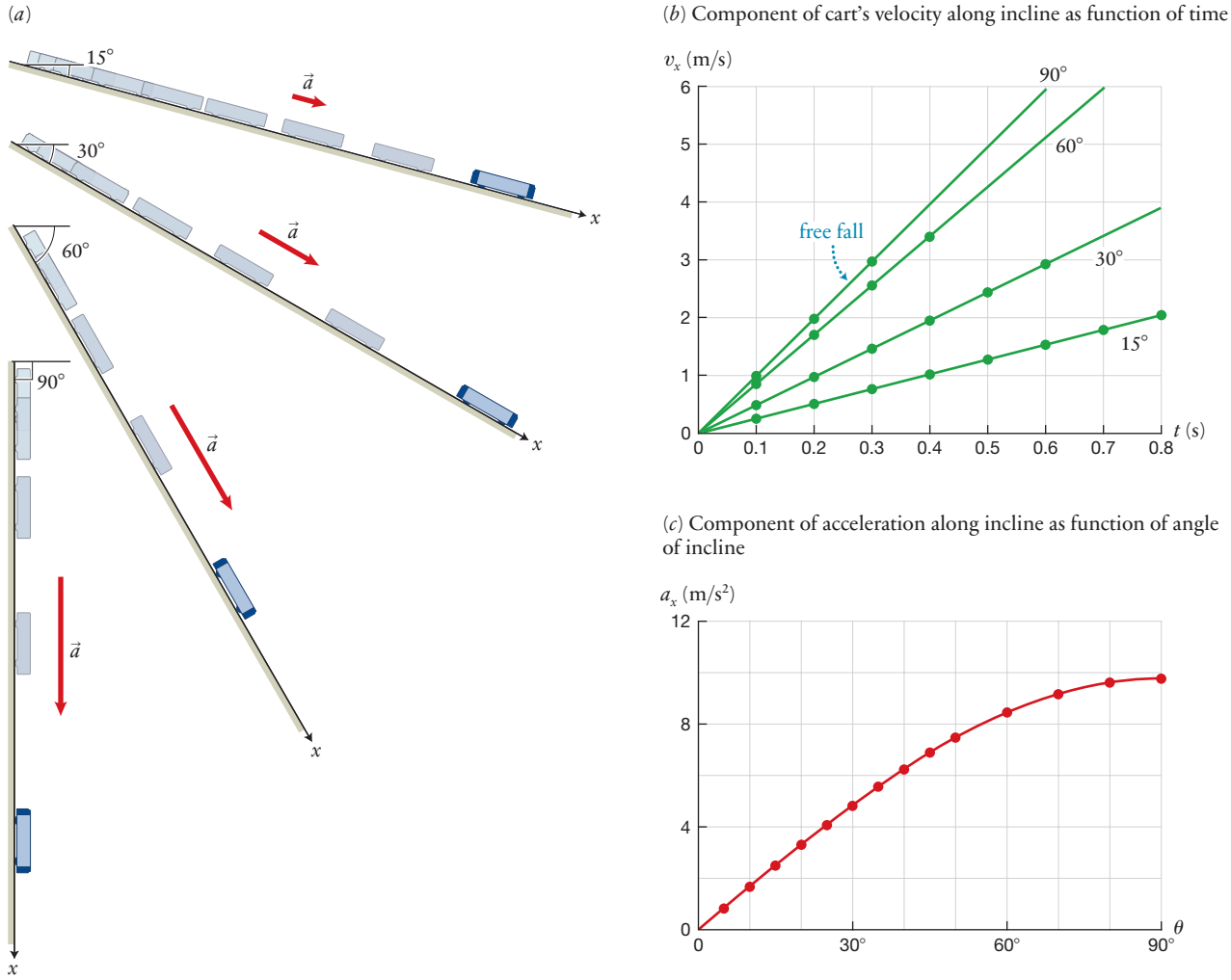
$$\frac{x_f}{t_f^2} = \tfrac{1}{2} a_x, \tag{3.19}$$

which tells us that Galileo's ratio (Eq. 3.17) is directly proportional to the x component of the acceleration. The fact that the ratio doesn't change as the ball rolls down the incline demonstrates that the acceleration is constant.

When Galileo increased the angle of the incline, the ratio x_f/t_f^2 increased, too, although for any fixed angle it remained constant regardless of the distance covered by the ball. Galileo was able to study only a limited range of angles, but he conjectured that for a given angle of incline the ratio x_f/t_f^2 would remain constant. For a vertical incline, he reasoned, the ball would be in free fall and still have a constant acceleration.

We saw in Section 3.2 that Galileo's conjecture was correct. **Figure 3.22*a*** shows a low-friction cart rolling down inclines of different angles. For each angle, the component of the velocity along the incline increases linearly with time, indicating that the component of the acceleration along the incline is constant for that angle. Moreover, the component of the acceleration along the incline (the slope of the lines in Figure 3.22*b*) becomes larger as the angle is increased. Figure 3.22*c* shows how the component of the acceleration along the incline varies with the angle of incline θ. Experimentally we

Figure 3.22 The acceleration of a low-friction cart rolling down an incline increases as the angle of incline increases.



determine that the component of the acceleration along the incline obeys the relationship

$$a_x = +g \sin \theta \tag{3.20}$$

when the x axis points downward along the incline. For a vertical incline, $\theta = 90°$ and Eq. 3.20 correctly yields the x component of the acceleration of a freely falling object.

3.13 As the angle θ of the incline used to collect the data of Figure 3.22 is increased beyond 90°, what happens to the acceleration? Does this result make sense (provided you always put the cart on the top side of the track)?

3.8 Instantaneous acceleration

So far, we have discussed only motion that has constant acceleration, but not all acceleration is constant. For example, when a car accelerates, the acceleration levels off as the car approaches top speed. This is illustrated in Figure 3.23. If the car's acceleration were constant, the $v_x(t)$ curve would

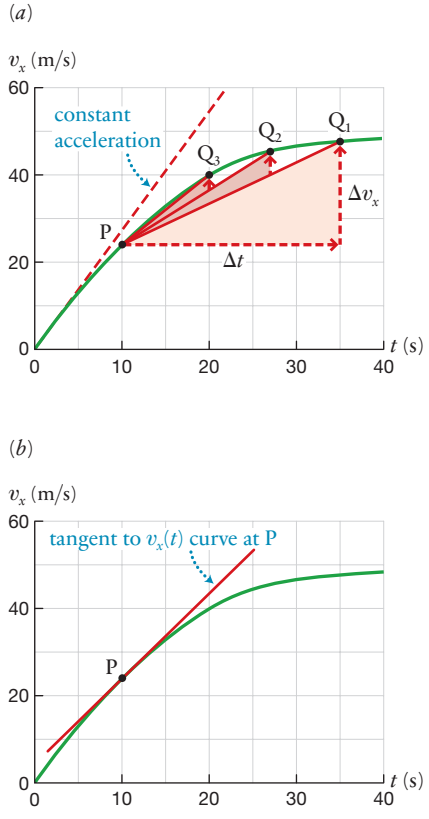


Figure 3.23 Velocity-versus-time graphs for an accelerating car. (*a*) As the time interval Δt over which the x component of the acceleration is averaged gets smaller, Δv_x gets smaller and the straight-line segment PQ approaches the tangent to the curve at P. (*b*) The slope of the tangent to the curve at P gives the x component of the instantaneous acceleration at that instant.

be the dashed straight line in Figure 3.23*a*; instead, the x component of the velocity levels off at about +50 m/s.

To determine the car's acceleration at a given instant, we follow the procedure of successive approximations we used in Section 2.9 to determine instantaneous velocity. Suppose we are interested in the x component of the car's acceleration 10 s after it begins accelerating. The x component of the *average* acceleration is obtained by dividing the change in the x component of the velocity during a certain time interval by the duration of that time interval (Eq. 3.1). For example, the x component of the car's average acceleration during the 25-s time interval following the instant we are interested in is given by the slope of the straight line segment PQ_1 in Figure 3.23*a*. To obtain the x component of the acceleration at $t = 10$ s, we make the time interval Δt over which we evaluate Δv_x successively shorter. As Δt approaches zero, the straight line segment PQ becomes tangent to the $v_x(t)$ curve at P, and the ratio $\Delta v_x/\Delta t$ approaches the x component of the **instantaneous acceleration** (or simply **acceleration**)* at point P:

$$a_x \equiv \lim_{\Delta t \to 0} \frac{\Delta v_x}{\Delta t}. \tag{3.21}$$

The right side of Eq. 3.21 is the definition of the derivative of v_x with respect to t, and so

$$a_x = \frac{dv_x}{dt}. \tag{3.22}$$

Note that the result is again a function of time. The acceleration-versus-time graph for the accelerating car of Figure 3.23 is shown in **Figure 3.24**. The slope of the $v_x(t)$ curve at a given point along the curve of Figure 3.23*b* gives the x component of the acceleration at that instant. For example, the value of the x component of the acceleration 10 s after the car begins accelerating is equal to the slope of the tangent to the $v_x(t)$ curve at point P (of Figure 3.23*b*).

In Section 2.9 we saw that the slope of the $x(t)$ curve for an object gives the x component of the object's velocity. Now we see that the slope of the $v_x(t)$ curve gives the x component of the object's acceleration:

$$\begin{array}{lcccc} & \textit{slope of} & & \textit{slope of} & \\ & x(t)\ \textit{curve} & x \text{ component of} & v_x(t)\ \textit{curve} & x \text{ component of} \\ \text{position } x & \longrightarrow & \text{instantaneous} & \longrightarrow & \text{instantaneous} \\ & & \text{velocity } \vec{v} & & \text{acceleration } \vec{a}. \end{array}$$

Mathematically this can be expressed as

$$a_x = \frac{dv_x}{dt} = \frac{d}{dt}\left(\frac{dx}{dt}\right) \equiv \frac{d^2x}{dt^2}, \tag{3.23}$$

where the expression on the right side represents the *second derivative* (the "derivative of the derivative") of x with respect to t.

Equation 3.23 allows us to obtain the x component of the acceleration of an object from its position. Often we are interested in the inverse procedure, however: What is the displacement during a given time interval for, say, the varying acceleration shown in Figure 3.24? The trick is to apply

*As with velocity, the word *instantaneous* is implied whenever we say "acceleration." Only when we want to speak of *average* acceleration will we use the adjective.

what we know about constant-acceleration motion over time intervals small enough to make the acceleration approximately constant over that time interval.

Suppose we want to know the displacement during the time interval between t_i and t_f. We first obtain the change in velocity during that time interval and then calculate the displacement.

We begin by dividing up the time interval between t_i and t_f into a number of equal shorter intervals δt and approximating the acceleration by the stepped series of short segments of constant acceleration as shown in Figure 3.24*a*. Applying Eq. 3.3 to each shorter interval δt, we can write the x component of the change in velocity during each segment as

$$(\Delta v_x)_n = a_x(t_n)\delta t, \tag{3.24}$$

where $n = 1, 2, \ldots, 7$ and $a_x(t_n)$ is the x component of the acceleration at t_n (this instant is at the middle of each short interval shown in Figure 3.24*a*). The product $a_x(t_n)\delta t$ is the area of the dark shaded rectangle in Figure 3.24*b*. The x component of the velocity change during the entire interval from t_i to t_f is then approximately equal to the sum of the $(\Delta v_x)_n$ for all the short intervals δt between t_i and t_f:

$$\Delta v_x \approx \sum_n (\Delta v_x)_n = \sum_n a_x(t_n)\delta t, \tag{3.25}$$

where $\sum a_x(t_n)\delta t$ represents the sum of all the rectangular areas under the steps between t_i and t_f (the light shaded area in Figure 3.24*b*).

To get a better approximation, we can divide the time interval $t_f - t_i$ into a larger number of smaller intervals δt, so that the variation of the acceleration over each interval becomes even smaller and the red stepped curve is closer to the actual (light red) $a_x(t)$ curve.

An exact result is obtained if we let δt approach zero (and the number of small intervals approaches infinity), as in Figure 3.24*c*:

$$\Delta v_x = \lim_{\delta t \to 0} \sum_n a_x(t_n)\delta t, \tag{3.26}$$

which is precisely the definition of the *integral* of the x component of the acceleration with respect to t from t_i to t_f:

$$\Delta v_x = \int_{t_i}^{t_f} a_x(t)dt. \tag{3.27}$$

This integral represents the area under the $a_x(t)$ curve between t_i and t_f in Figure 3.24*c*. Integrals, like derivatives, can be evaluated mathematically, providing us with a convenient method for calculating the right side of Eq. 3.27 and allowing us to determine velocity from acceleration.

Once we know the velocity, we can use the same approach to obtain the displacement from the velocity. **Figure 3.25** graphically shows how to determine the x component of the displacement by evaluating the area under the $v_x(t)$ curve. In analogy to Eq. 3.27 we can thus write

$$\Delta x = \int_{t_i}^{t_f} v_x(t)dt. \tag{3.28}$$

This expression allows us to mathematically determine the x component of the displacement from a $v_x(t)$ curve. The x component of the velocity, in turn, can be obtained from the $a_x(t)$ curve (Eq. 3.27).

Figure 3.24 Determining the change in the x component of the velocity for an accelerating object.

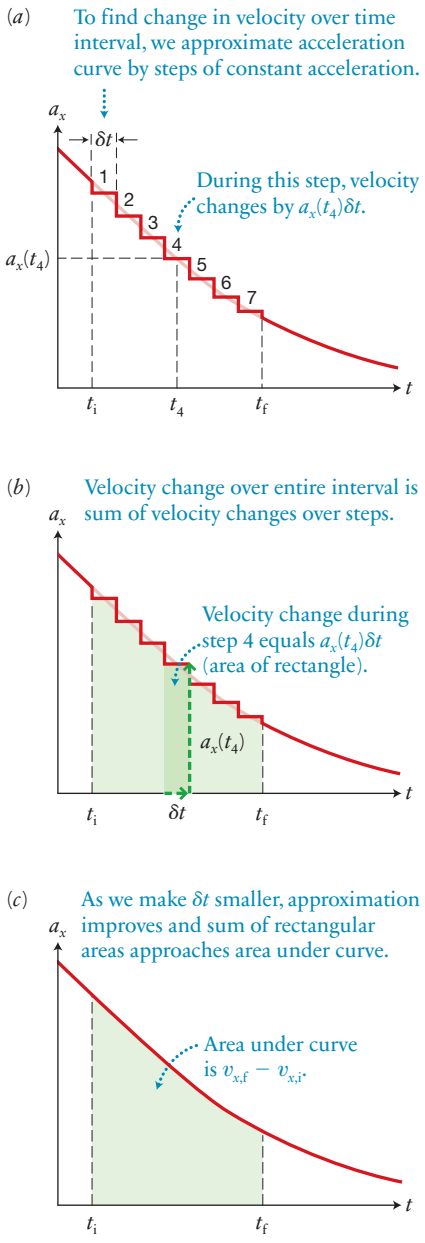
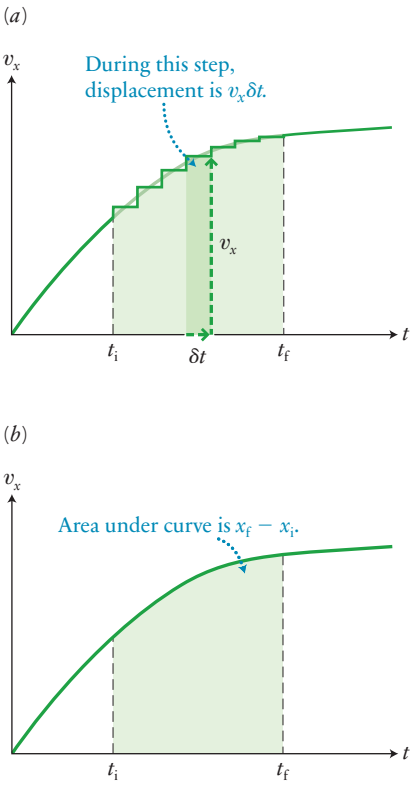


Figure 3.25 (*a*) To determine the *x* component of the displacement, we approximate the $v_x(t)$ curve by short segments of constant velocity. (*b*) When the segments are made infinitesimally small, the *x* component of the displacement becomes equal to the area under the $v_x(t)$ curve.



In Section 2.8, I showed that for motion at constant velocity the area under a $v_x(t)$ curve is equal to the *x* component of the object's displacement. Now we see that this holds true for accelerated motion as well. The geometrical relationship among acceleration, velocity change, and displacement is thus

acceleration a_x $\xrightarrow{\text{area under } a_x(t) \text{ curve}}$ velocity change Δv_x from t_i *to* t_f $\xrightarrow{\text{area under } v_x(t) \text{ curve}}$ displacement Δx from t_i *to* t_f.

EXERCISE 3.7 Using calculus to determine displacement

Suppose an object initially at x_i at $t_i = 0$ has a constant acceleration whose *x* component is a_x. Use calculus to show that the *x* component of the velocity and the *x* coordinate at some final instant t_f are given by Eq. 3.10 and Eq. 3.9, respectively.

SOLUTION Because the acceleration is constant, I can pull a_x out of the integration in Eq. 3.27:

$$\Delta v_x = v_{x,f} - v_{x,i} = \int_{t_i}^{t_f} a_x \, dt = a_x \int_{t_i}^{t_f} dt = a_x(t_f - t_i).$$

Substituting $t_i = 0$ and rearranging terms, I obtain Eq. 3.10:

$$v_{x,f} = v_{x,i} + a_x t_f. \checkmark$$

For an arbitrary final instant *t*, I can drop the subscript f. Substituting this expression into Eq. 3.28, I get

$$\Delta x = \int_{t_i}^{t_f} (v_{x,i} + a_x t) dt$$

or, pulling constant terms out of the integration and then carrying out the integration,

$$\Delta x = x_f - x_i = v_{x,i} \int_{t_i}^{t_f} dt + a_x \int_{t_i}^{t_f} t \, dt = v_{x,i} t_f + a_x \left[\tfrac{1}{2} t^2 \right]_{t_i}^{t_f},$$

which yields Eq. 3.9:

$$x_f = x_i + v_{x,i} t_f + \tfrac{1}{2} a_x t_f^2. \checkmark$$

3.14 Take the first and second time derivatives of x_f in Eq. 3.11. What do you notice?

Chapter Summary

Accelerated motion (Sections 3.1, 3.4, 3.5, 3.8)

Concepts If the velocity of an object is changing, the object is **accelerating.** The x component of an object's **average acceleration** is the change in the x component of its velocity divided by the time interval during which this change takes place.

The x component of the object's **instantaneous acceleration** is the x component of its acceleration at any given instant.

A **motion diagram,** which is constructed as described in the Procedure box on page 93, shows the positions of a moving object at equally spaced time intervals.

Quantitative Tools The x component of the **average acceleration** is

$$a_{x,\text{av}} \equiv \frac{\Delta v_x}{\Delta t} = \frac{v_{x,\text{f}} - v_{x,\text{i}}}{t_\text{f} - t_\text{i}}. \tag{3.1}$$

The x component of the **instantaneous acceleration** is

$$a_x = \frac{dv_x}{dt} = \frac{d^2x}{dt^2}. \tag{3.23}$$

The x component of the change in velocity over a time interval is given by

$$\Delta v_x = \int_{t_\text{i}}^{t_\text{f}} a_x(t)dt. \tag{3.27}$$

The x component of the displacement over a time interval is given by

$$\Delta x = \int_{t_\text{i}}^{t_\text{f}} v_x(t)dt. \tag{3.28}$$

Motion with constant acceleration (Section 3.5)

Concepts If an object has constant acceleration, the $v_x(t)$ curve is a straight line that has a nonzero slope and the $a_x(t)$ curve is a horizontal line.

Quantitative Tools If an object moves in the x direction with constant acceleration a_x starting at $t = 0$, with initial velocity $v_{x,\text{i}}$ at initial position x_i, its x coordinate at any instant t is given by

$$x(t) = x_\text{i} + v_{x,\text{i}}t + \tfrac{1}{2}a_x t^2. \tag{3.11}$$

The x component of its instantaneous velocity is given by

$$v_x(t) = v_{x,\text{i}} + a_x t, \tag{3.12}$$

and the x component of its final velocity is given by

$$v_{x,\text{f}}^2 = v_{x,\text{i}}^2 + 2a_x\Delta x. \tag{3.13}$$

Free fall and projectile motion (Sections 3.2, 3.3, 3.6)

Concepts An object subject only to gravity is in **free fall.** All objects in free fall near the surface of Earth have the same acceleration, which is directed downward. We call this acceleration the **acceleration due to gravity** and denote its magnitude by the letter g.

An object that is launched but not self-propelled is in **projectile motion.** Once it is launched, it is in free fall. The path it follows is called its **trajectory.**

Quantitative Tools The magnitude g of the downward acceleration due to gravity is

$$g = |\vec{a}_{\text{free fall}}| = 9.8\ \text{m/s}^2 \quad (\text{near Earth's surface}). \tag{3.14}$$

Motion on an inclined plane (Section 3.7)

Concepts An object moving up or down an inclined plane on which friction is negligible has a constant acceleration that is directed parallel to the surface of the plane and points downward along the surface.

Quantitative Tools When friction is negligible, the x component of acceleration a_x for an object moving on an inclined plane that rises at an angle θ above the horizontal is

$$a_x = +g\sin\theta \tag{3.20}$$

when the x axis is directed downward along the plane.

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = easy, •• = intermediate, ••• = hard; CR = context-rich problem.

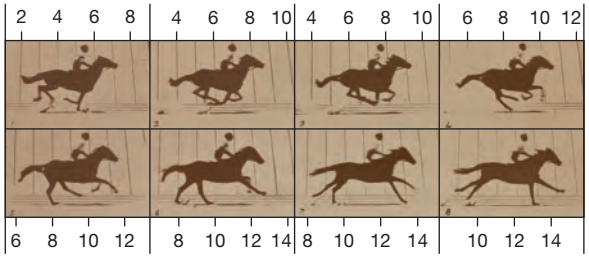
3.1 Changes in velocity

1. Figure P3.1 is based on a multiple-flash photographic sequence of an object moving from left to right on a track, as seen from above. The time intervals between successive flashes are all the same. During which portion(s) of the motion is the object (*a*) speeding up and (*b*) slowing down? Explain how you know in each case. (*c*) How would your answers change if the object were moving from right to left? •

Figure P3.1

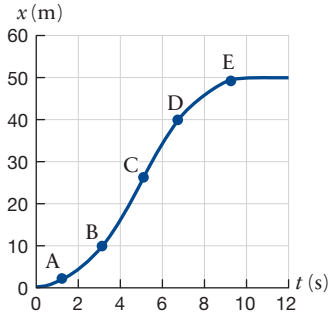
2. A bus is travelling in the northeast direction. What are the directions of its acceleration and velocity (*a*) if it is speeding up and (*b*) if it is slowing down? •
3. Figure P3.3 shows a series of photographs of a racehorse taken in 1877 by Eadweard Muybridge (1830–1904). Muybridge used multiple, equally spaced cameras triggered sequentially at equal time intervals. Is the horse accelerating? How can you tell? •

Figure P3.3



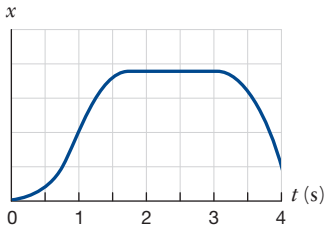
4. Figure P3.4 shows a graph of position as a function of time for an object moving along a horizontal surface. At which of the labeled points is the object speeding up? •

Figure P3.4



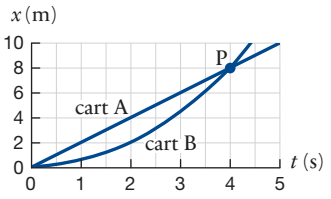
5. While you are driving on a highway, a friend in another car passes you at a steady speed. You accelerate at a constant rate to catch up. When you do catch up, are your two cars going the same speed? Use a graph to support your answer. ••
6. The position of a person pacing in a hall is given by the $x(t)$ curve in Figure P3.6. During which time interval(s) is the acceleration (*a*) positive and (*b*) negative? (*c*) Is the acceleration ever zero during the 4-s interval shown? ••

Figure P3.6



7. You are travelling by a bullet train from a point to a distance at 500 km. If the bullet train operates at a uniform speed of 200 km/h, describe your acceleration and inertia at different stages during this trip. ••
8. Figure P3.8 shows the position curves for two carts, A and B, moving on parallel tracks along a horizontal surface. The instant when the two carts are at the same distance from their starting point is indicated by point P. Which cart has the greater acceleration at that instant? ••

Figure P3.8



9. Two cars are headed south on the highway at different speeds. Starting when the faster car pulls alongside the slower car, one of the cars accelerates northward for 5.0 s and the other accelerates southward for 5.0 s. At the end of that 5.0-s interval, the two cars have the same speed. Which car is now ahead? •••

3.2 Acceleration due to gravity

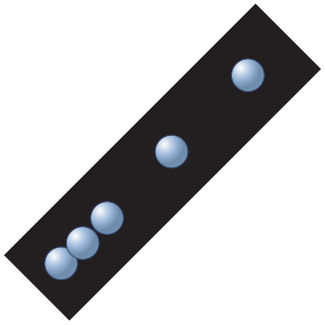
10. A tennis ball was dropped from rest from the window of a multistoried building. Calculate the distance travelled by the ball during one second considering the acceleration due to gravity (g) = 9.8 m/s^2. •
11. (*a*) What is the average speed, over the first 1.0 s of its motion, of a pebble released from rest off a bridge? (*b*) What is the pebble's average speed over the second 1.0 s of its motion? (*c*) What is its average speed during these first 2.0 s of its motion? •

12. You toss a ball to your brother leaning out of a window h m above you, throwing just hard enough for it to reach him. At the same instant, he drops a chocolate bar to you. Prove that the chocolate bar and the ball pass each other $h/4$ m above the ground. ••
13. Suppose that the acceleration due to gravity near Earth's surface was cut in half (to about 5 m/s^2). How would this affect the graphs shown in Figure 3.6, assuming that everything else about the experiment remains the same? Sketch the new curves. ••

3.3 Projectile motion

14. A photographer shows you a multiple-flash photographic sequence of a ball traveling vertically (Figure P3.14). (*a*) If the ball is traveling downward, what is the correct orientation of the picture? (*b*) If the ball is traveling upward, what is the correct orientation of the picture? (*c*) According to the evidence provided by the picture, what is the direction of the ball's acceleration in parts *a* and *b*? •

Figure P3.14

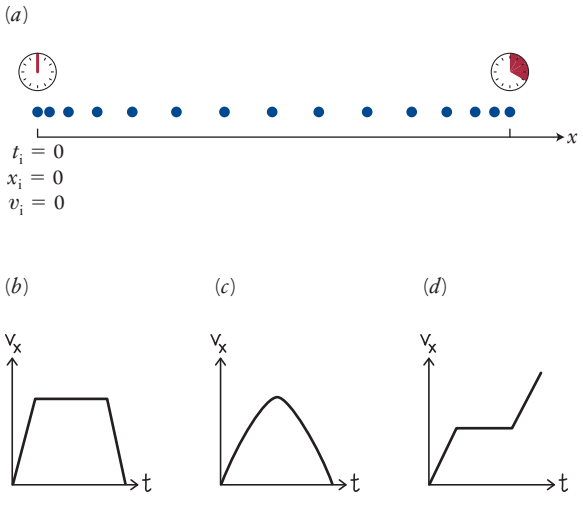


15. A grenade launcher shoots a grenade vertically upward at an initial speed of 120 m/s. What are the vertical distances covered by it after (*a*) 6.0 s, (*b*) 12.0 s? •
16. Two identical balls have the same inertia. One was dropped from a building 20 m high, and another was thrown up from the ground at an initial speed of 26 m/s. After how long will both the balls meet each other in air? •
17. A ball is shot straight up near the Earth's surface. (*a*) Can its velocity be zero while its acceleration is nonzero? (*b*) Is the velocity of the ball the same at any two points, one of which is on the upward trajectory, and one of which is on the downward one? (*c*) Is the speed of the ball the same at any two points, one of which is on the upward trajectory, and one of which is on the downward one? ••
18. How would graphs *c* and *d* in Figure 3.8 be different if the same experiment was conducted on the surface of the Moon, where the acceleration due to gravity is six times less than it is on Earth? ••
19. You throw snowballs down to the sidewalk from the roof of a building. Which technique, if either, makes the snowballs land with more speed: throwing them straight down as hard as you can or throwing them straight up just as hard? (Ignore air resistance.) ••
20. You are standing by a window and see a ball, thrown from below, moving up past the window. The ball is visible for a time interval Δt_{up}. On its way back down the ball passes the window again, remaining visible for a time interval Δt_{down}. Neglecting the effects of air resistance, is $\Delta t_{up} > \Delta t_{down}$, $\Delta t_{up} = \Delta t_{down}$, $\Delta t_{up} < \Delta t_{down}$, or is it impossible to tell without more information? Briefly explain your reasoning, using graphs. ••
21. For an object in free fall, the curve of its velocity as a function of time is a straight line. How would this curve be different when air resistance (which changes with speed) is not negligible? •••
22. When air resistance is ignored, it is straightforward to calculate the travel time interval for a Ping-Pong ball tossed up into the air and caught on its way down because the acceleration can be taken to be a constant 9.8 m/s^2 downward. However, a more accurate result is obtained when air resistance is taken into account. When this is done, is the magnitude of the ball's acceleration greater than, less than, or the same as when air resistance is ignored (*a*) as the ball rises and (*b*) as it falls? (*c*) Do you expect the travel time interval obtained in the more accurate analysis to be greater than, less than, or the same as in the simplified case? •••

3.4 Motion diagrams

23. Draw a motion diagram for a car that starts from rest and accelerates at 4.0 m/s^2 for 10 s. •
24. A car accelerating from rest at constant acceleration reaches a speed of 30 km/h in 5.0 s. Draw a motion diagram for the car. •
25. The motion of a cart moving along a horizontal surface is described by the motion diagram shown in Figure P3.25*a*. The position of the cart is measured every 0.5 s. Asked to suggest a qualitative graph of velocity versus time that would correspond to this motion, three of your classmates draw the graphs shown in Figures P3.25*b*–*d*. Which graph is correct? •

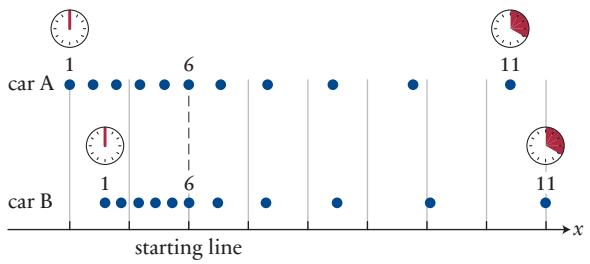
Figure P3.25



26. Draw a motion diagram for a car that starts from rest and accelerates at 5.0 m/s^2 for 6.0 s, then travels with constant speed for 10 s, and then slows to a stop with constant acceleration in 4.0 s. ••

27. A ball you throw straight up has a speed of 30 m/s when it leaves your hand. Draw a motion diagram for the ball up to the instant at which it has a speed of 30 m/s again. Indicate the instant at which the ball reaches its highest position. ●●
28. Figure P3.28 shows motion diagrams for two cars, A and B, beginning a race. The diagrams show the position of each car at instants separated by equal time intervals. Both cars drive up to the starting line and then begin to accelerate. (*a*) Which car has the greater velocity at instant 6? How can you tell? (*b*) Which car has the greater velocity at instant 11? How can you tell? (*c*) Which car has the greater acceleration after instant 6? How can you tell? ●●

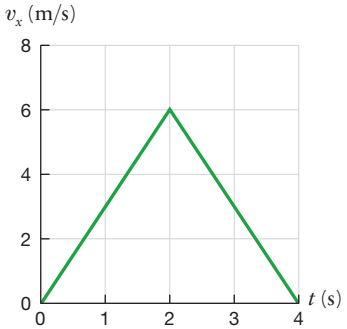
Figure P3.28



3.5 Motion with constant acceleration

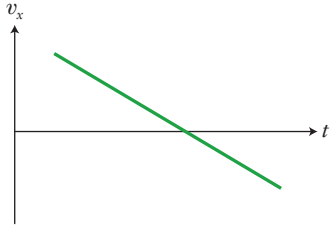
29. A sports car goes from zero to 108 km/hr in 3 seconds. (*a*) What is the acceleration of the car, assuming it is constant? (*b*) How far does it travel during this time period? ●
30. You and your little brother are rolling toy cars back and forth to each other across the floor. He is sitting at $x = 0$, and you are at $x = 4.0$ m. You roll a car toward him, giving it an initial speed of 2.5 m/s. It stops just as it reaches him in 3.0 s. (*a*) In what direction is the car's acceleration? (*b*) What is the car's average acceleration? ●
31. An electron is decelerated from 4×10^6 m/s to zero in 6×10^{-8} s. (*a*) Calculate the distance travelled by the electron in this interval. (*b*) Calculate the average acceleration. ●
32. Figure P3.32 shows a graph of velocity as a function of time for a cart moving along a horizontal surface. When asked to describe the motion that resulted in this graph, a student states, "The cart first moved forward with a constant speed, reached a maximum distance at time 2 seconds, then turned around and came back to its initial position at time 4 seconds." What is your assessment of this comment? ●

Figure P3.32



33. (*a*) How do you determine an object's displacement from a velocity-versus-time graph? (*b*) What distance does a car travel as its speed changes from 0 to 20 m/s in 10 s at constant acceleration? (*c*) The x component of the average velocity of a particle moving along the x axis at constant acceleration is $v_{x,\text{av}} = \frac{1}{2}(v_{x,\text{i}} + v_{x,\text{f}})$. Is the x component of the average velocity also given by this expression when the acceleration is not constant? Use a velocity-versus-time graph to support your answer. ●●
34. On a freeway entrance ramp, you accelerate your 1200-kg car from 5 m/s to 20 m/s over 500 m. Having an open lane, you continue to accelerate at the same (constant) rate for another 500 m. Is your final speed 35 m/s, less than 35 m/s, or greater than 35 m/s? ●●
35. In an introductory physics laboratory, a student drops a steel ball of radius 15 mm, and a device records its position as a function of time. The clock on the device is set so that $t = 0$ at the instant the ball is dropped. What is the ball's displacement (*a*) between 0.15 s and 0.25 s and (*b*) between 0.175 s and 0.275 s? ●●
36. Compare the acceleration of a car that goes from 0 to 100 km/hr in 3 seconds to that of a boulder falling freely near the surface of the Earth. ●●
37. An electron in an old-fashioned television picture tube is accelerated, at a constant rate, from 2.0×10^5 m/s to 1.0×10^7 m/s in a 12-mm-long "electron gun." (*a*) What is the acceleration of the electron? (*b*) What time interval is needed for the electron to travel the length of the gun? ●●
38. In a car moving at constant acceleration, you travel 250 m between the instants at which the speedometer reads 40 km/h and 60 km/h. (*a*) How many seconds does it take you to travel the 250 m? (*b*) What is your acceleration? ●●
39. Based on the $v(t)$ curve in Figure P3.39, explain whether each of these statements is necessarily true for the time interval shown: (*a*) The acceleration is constant. (*b*) The object passes through the position $x = 0$. (*c*) The object has zero velocity at some instant. (*d*) The object is always moving in the same direction. ●●

Figure P3.39



40. In October 1997, Andy Green broke the sound barrier on land in a jet-powered car traveling at 763 mi/h over a 1.0-mi course. As it arrived at the beginning of this measured mile, Green's car had accelerated from rest to 763 mi/h over a 5.0-mi stretch. (*a*) What was Green's average acceleration in that 5.0 mi? (*b*) Calculate the time interval needed for Green to reach his maximum speed. ●●

41. A driver starts a car starts from rest and accelerates to 72 km/hr in 10 seconds. He drives at this speed for another 20 s, and then slows to a stop over a distance of 500 m. (*a*) What is the acceleration of the car during the first and last legs of this trip? (*b*) What is the total travel time? (*c*) What is the total distance of the trip? ••
42. In rush-hour traffic, the car in front of you suddenly puts on the brakes. You apply your brakes 0.50 s later. The accelerations of the two cars are the same. Does the distance between the two cars remain constant, decrease, or increase? ••
43. Two cars are moving at 97 km/h, one behind the other, on a rural road. A deer jumps in front of the lead car, and its driver slams on the brakes and stops. What minimum initial distance between the rear of the lead car and the front of the second car is required if the second car is to stop before hitting the lead car? Assume that the acceleration is the same for both cars and that the driver of the second car begins braking 0.50 s after the lead car begins braking. •••
44. As a physics instructor hurries to the bus stop, her bus passes her, stops ahead, and begins loading passengers. She runs at 6.0 m/s to catch the bus, but the door closes when she's still 8.0 m behind the door, and the bus leaves the stop at a constant acceleration of 2.0 m/s^2. She has missed her bus, but as a physics exercise she keeps running at 6.0 m/s until she draws even with the bus door. (*a*) Calculate Δt from the instant the bus leaves the stop to the instant the instructor draws even with the door. (*b*) At the instant she draws even, how fast is she moving? (*c*) At the instant she draws even, how fast is the bus moving? (*d*) Plot $v(t)$ for the bus and the instructor. (*e*) Plot $x(t)$ for the bus and the instructor. •••
45. The day after the incident described in Problem 44, the instructor finds herself in the same situation. This time, she tries a harder physics exercise. She keeps running at a constant 6.0 m/s after drawing even with the bus door and pulls ahead for a while, but the accelerating bus soon overtakes her. By what maximum distance does she get ahead of the door? •••

3.6 Free-fall equations

46. A ball was dropped from the top of the Eiffel Tower, 300 m high. (*a*) Calculate the time taken by the ball to reach the ground (*b*) What is the final speed of the ball when it hits the ground? •
47. An object takes 3 s to fall from the top of the building to the ground. Calculate what time will be required by the object to fall from the same height on Moon's surface ($g_{moon} = 1.625$ m/s^2) •
48. A mortar-style fireworks shell is launched upward at 35 m/s. Draw a motion diagram showing the position of the shell and its velocity vector at $t = 0$, 1 s, 2 s, 3 s, and 4 s, and draw a graph of $x(t)$. •
49. The distance a ball thrown upward and released at $t = 0$ travels between $t = 3.0$ s and $t = 4.0$ s is the same as the distance the ball travels between $t = 2.0$ s and $t = 3.0$ s. How is this possible? ••
50. You toss a ball straight up, and it reaches a maximum height h above the launch position. If you want to double the length of the time interval during which the ball stays in the air above the launch position, to what maximum height do you have to throw it? ••
51. A ball is thrown vertically up with a speed 10 m/s, caught by a boy leaning out of a window in 4 seconds. (*a*) Calculate the height of the window from the ground (*b*) What is the final speed of the ball before it is caught? ••
52. A snowshoer falls off a ridge into a snow bank 3.6 m below and penetrates 0.80 m into the snow before stopping. What is her average acceleration in the snow bank? ••
53. You are on the second floor of a building under construction, laying bricks of dimensions 63 mm × 89 mm × 170 mm. At your command, a coworker below tosses bricks up to you, and you catch them when they have risen 5.0 m above the launch point. He tosses the bricks consistently so that, if untouched by you, they would rise to a maximum height of 6.0 m above the launch point. (*a*) Soon you both get good at this, and you know exactly how many seconds after you yell "3, 2, 1, throw!" you should reach out and catch the brick on its way up. What is the length of this time interval if your coworker throws without delay at your command? (*b*) If you miss a brick on the way up, during what remaining time interval can you reach out and catch the brick on its way down? ••
54. A rocket ignited on the ground travels vertically upward with an acceleration of magnitude $4g$. A spent rocket stage detaches from the payload after the rocket has accelerated for 5.0 s. With what speed does the spent stage hit the ground? ••
55. Show that the distances traveled by a falling rock in successive 1-s intervals after release are in the ratio 1:3:5:7. . . . ••
56. A hot-air balloon takes off from the ground traveling vertically with a constant upward acceleration of magnitude $g/4$. After time interval Δt, a crew member releases a ballast sandbag from the basket attached to the balloon. How many seconds does it take the sandbag to reach the ground? •••
57. A rock dropped from the top of a building takes 0.50 s to fall the last 50% of the distance from the top to the ground. How tall is the building? •••
58. A hot-air balloon of diameter 10 m rises vertically at a constant speed of 12 m/s. A passenger accidentally drops his camera from the railing of the basket when it is 18 m above the ground. If the balloon continues to rise at the same speed, how high is the railing when the camera hits the ground? •••
59. A ball tossed vertically upward from the ground next to a building passes the bottom of a window 1.8 s after being tossed and passes the top of the window 0.20 s later. The window is 2.0 m high from top to bottom. (*a*) What was the ball's initial velocity? (*b*) How far is the bottom of the window from the launch position? (*c*) How high does the ball rise above the launch position? •••

3.7 Inclined planes

60. In a children's park, a boy uses a slide with inclination 50° with respect to the ground. Ignoring friction, calculate the magnitude of the acceleration of the boy. •
61. A box slides from rest down a 2-m long 20° incline. How long does it take the box to reach the bottom of the incline? •

62. Imagine that in his experiment with balls rolling down an inclined plane, Galileo had considered balls that were given an initial speed. Would he have reached different conclusions about the ratio of distance and square of time? •

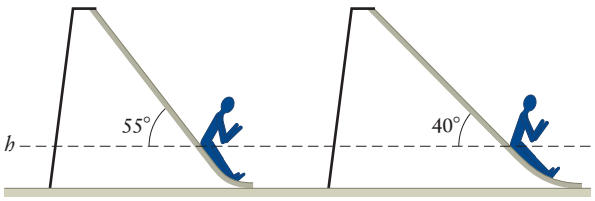
63. A worker releases boxes at the top of a ramp. From the bottom of the ramp the boxes slide 10 m across the floor to a barrier wall. If the ramp is at an elevation angle of 20°, how long must the ramp be if the boxes are to reach the wall 2.0 s after leaving the ramp? Ignore friction. ••

64. A man steps outside one winter day to go to work. His icy driveway is 8.0 m long from top to mailbox, and it slopes downward at 20° from the horizontal. He sets his briefcase on the ice at the top while opening the garage, and the briefcase slides down the driveway. Ignore friction. (*a*) What is its acceleration? (*b*) How many seconds does it take to get halfway to the mailbox? (*c*) How many seconds does it take to reach the mailbox? (*d*) What is its speed at the instant it reaches the mailbox? ••

65. You and a friend ride what are billed as the "world's longest slides" at a county fair. Your slide is 100 m long, and your trip takes 10 s, including any effect of friction. Your friend chooses a taller, 150-m-long slide made from the same material as yours and with the same angle of incline. Ignore friction. (*a*) What is the magnitude of your acceleration down the slide? (*b*) What is the magnitude of your friend's acceleration down the slide? (*c*) How many seconds does it take your friend to get to the bottom? (*d*) What is your speed when you hit the bottom? (*e*) What is your friend's speed when she reaches the bottom? ••

66. Two children at a playground slide from rest down slides that are of equal height but are inclined at different angles with respect to the horizontal (Figure P3.66). Ignoring friction, at height h above the ground, which child has (*a*) the greater acceleration and (*b*) the higher speed? ••

Figure P3.66



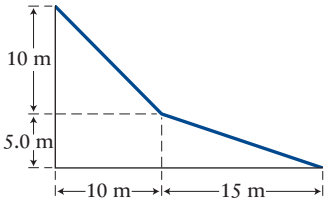
67. You are playing air hockey with a friend. The puck is sitting at rest in his goal when he suddenly lifts his end of the table by 0.50 m. The puck slides down the tilted surface into your goal, 2.4 m away. Ignore friction. (*a*) How many seconds does it take the puck to reach your goal? (*b*) With what speed does the puck hit your goal? ••

68. A block has an initial speed of 6.0 m/s up an inclined plane that makes an angle of 37° with the horizontal. Ignoring friction, what is the block's speed after it has traveled 2.0 m? ••

69. A box is at the lower end of a very slippery ramp of length ℓ that makes a nonzero angle θ with the horizontal. A worker wants to give the box a quick shove so that it reaches the top of the ramp. (*a*) How fast must the box be going after the shove for it to reach its goal? (Ignore the distance traveled while the shove is executed.) (*b*) What is its speed halfway up the ramp? ••

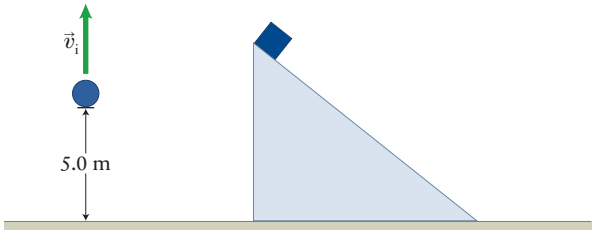
70. A skier is at the top of a run that consists of two slopes having different inclines (Figure P3.70). The skier lets go, beginning the run with essentially zero speed. Ignoring friction, what are (*a*) his speed at the end of the lower slope and (*b*) his average acceleration over the entire run? ••

Figure P3.70



71. A ball is projected vertically upward from an initial position 5.0 m above the ground (Figure P3.71). At the same instant, a cube is released from rest down an ice-covered incline, from a height not necessarily equal to 5.0 m. The two objects reach the ground at the same instant, and both have a final speed of 15 m/s. What is the angle of the incline? •••

Figure P3.71



72. You hold a puck at the top of an ice-covered ramp inclined at 60° with respect to the vertical. Your friend stands nearby on level ground and holds a ball at the same height h above ground as the puck. If the puck and the ball are released from rest at the same instant, what is the ratio of the time interval it takes the ball to reach the ground to the time interval it takes the puck to reach the end of the ramp? •••

73. A child on a sled slides down an icy slope, starting at a speed of 2.5 m/s. The slope makes a 15° angle with the horizontal. After sliding 10 m down the slope, the child enters a flat, slushy region, in which she slides for 2.0 s with a constant negative acceleration of $-1.5\ \text{m/s}^2$ with respect to her direction of motion. She then slides up another icy slope that makes a 20° angle with the horizontal. How far up the second slope does she slide? •••

3.8 Instantaneous acceleration

74. You throw a ball straight up with an initial speed of 10 m/s. (*a*) What is the ball's instantaneous acceleration at instant t_1, just after it leaves your hand; at instant t_2, the top of its trajectory; and at instant t_3, just before it hits the ground? (*b*) What is its average acceleration for the upward portion of its journey (t_1 to t_2) and for the entire trip (t_1 to t_3)? •

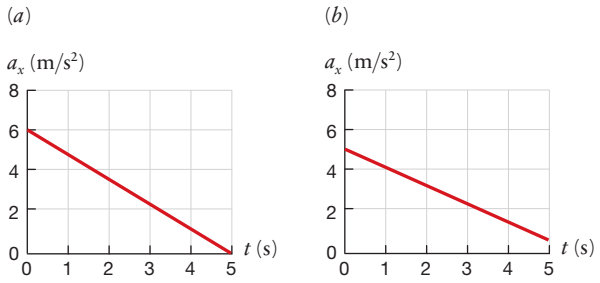
75. The equation of motion of a particle on xy plane as a function of time is given by

$$\vec{x} = 4bt^3\hat{i} + 48t + 20$$

Compute the particle's acceleration as a function time at 3 s. Here, $b = 4.5\ \text{m/s}^3$ •

76. Figure P3.76 shows graphs of the x component of acceleration as a function of time for two different carts rolling along a flat horizontal table. In which case is the change in the x component of velocity greater over the time interval shown? •

Figure P3.76



77. The position of a cart on a low-friction track can be represented by the equation $x(t) = b + ct + et^2$, where $b = 4.00$ m, $c = 6.00$ m/s, and $e = 0.200\ \text{m/s}^2$. (*a*) Is the cart accelerating? If so, is the acceleration constant? (*b*) What is the average velocity between $t_i = 0.200$ s and $t_f = 0.400$ s? (*c*) What is the velocity at $t = 0.200$ s and at $t = 0.400$ s? (*d*) What is the average acceleration between $t_i = 0.200$ s and $t_f = 0.400$ s? (*e*) What is the acceleration at $t = 0.200$ s and at $t = 0.400$ s? ••

78. A particle moves in the x direction according to the equation $x(t) = a + bt + ct^2$, where $a = 2.0$ m, $b = 0.5$ m/s, and $c = -0.4\ \text{m/s}^2$. ••

(*a*) What is the particle's instantaneous velocity and instantaneous acceleration at $t = 10$ s?

(*b*) What is its average velocity and average acceleration between $t = 5$ s and $t = 10$ s?

79. The acceleration of a particle can be represented by the equation $a(t) = t^2$. The particle begins accelerating from rest at $t = 0$.

(*a*) What is its acceleration at $t = 2$ s?

(*b*) What is the velocity at $t = 2$ s?

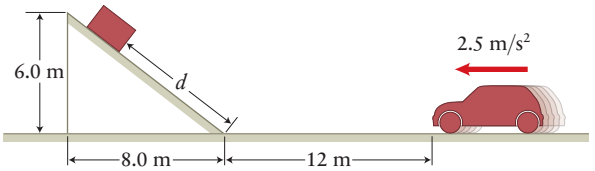
(*c*) What is the average acceleration between $t = 0$ and $t = 2$ s?

(*d*) What is the distance traveled between $t = 0$ and $t = 2$ s? ••

80. The acceleration of a particular car during braking has magnitude bt, where t is the time in seconds from the instant the car begins braking, and $b = 2.0\ \text{m/s}^3$. If the car has an initial speed of 50 m/s, how far does it travel before it stops? ••

81. A car is 12 m from the bottom of a ramp that is 8.0 m long at its base and 6.0 m high (Figure P3.81). The car moves from rest toward the ramp with an acceleration of magnitude $2.5\ \text{m/s}^2$. At some instant after the car begins moving, a crate is released from rest from some position along the ramp. The crate and car reach the bottom of the ramp at the same instant and at the same speed. (*a*) At what distance d up the ramp was the crate released? (*b*) How many seconds after the car started was the crate released? •••

Figure P3.81



82. In a laboratory experiment, a sphere of diameter 8.0 mm is released from rest at $t = 0$ at the surface of honey in a jar, and the sphere's downward speed v when it travels in the honey is found to be given by $v = v_{max}(1 - e^{-t/\tau})$, where $v_{max} = 0.040$ m/s and $\tau = 0.50$ s. (*a*) Obtain an expression for $a(t)$. (*b*) Draw graphs for $v(t)$ and $a(t)$ for the time interval 0 to 2.0 s. (*c*) Obtain an expression for $x(t)$, choosing the positive x axis as downward, and draw the graph for this function. (*d*) Use your $x(t)$ graph to determine the time interval needed for the sphere to reach the bottom of the container if the surface of the honey is 0.10 m above the bottom of the jar. •••

83. A video recording of a moving object shows the object moving up and down with a velocity given by $v_x(t) = v_{max}\cos(\omega t)$, where $v_{max} = 1.20$ m/s and $\omega = 0.15\text{s}^{-1}$ and where the positive x direction is upward. (*a*) Derive an expression for $a(t)$, the object's acceleration as a function of time. (*b*) If the object's position at $t = 0$ is $x(0) = 0$, derive an expression for $x(t)$, its position as a function of time. •••

Additional Problems

84. The equation of motion of a particle travelling along the x-axis is given as

$$x(t) = 3bt^3 + 4ct^2 + dt$$

where $b = 0.4\ \text{m/s}^3$, $c = 0.5\ \text{m/s}^2$ and $d = 50$ m/s.

(*a*) Calculate the particle's acceleration $t = 20$ s.

(*b*) What is the distance travelled at $t = 5$ s? •

85. It takes you 7.0 m to brake to a panic stop from a speed of 9.0 m/s. Using the same acceleration, how far do you go as you brake to a panic stop from a speed of 27 m/s? •

86. A person standing on a building ledge throws a ball vertically from a launch position 45 m above the ground. If it takes 2.0 s for the ball to hit the ground, (*a*) with what initial speed was the ball thrown and (*b*) in which direction was it thrown? ••

87. A bullet fired straight through a board 0.10 m thick strikes the board with a speed of 480 m/s, has constant acceleration through the board, and emerges with a speed of 320 m/s. (*a*) What is the component of the acceleration in the direction of the bullet's motion? (*b*) What is the time interval between the instant the bullet strikes the front face of the board and the instant it exits the back face? (*c*) Assuming the average acceleration you calculated in part *a*, what minimum thickness must the board have in order to stop the bullet? ••

88. During a blackout, you are trapped in a tall building. You want to call rescuers on your cell phone, but you can't remember which floor you're on. You pry open the doors to the elevator shaft, drop your keys down the shaft, and hear them hit bottom at ground level 3.27 s later. (*a*) Making a simple calculation and remembering that the ground floor is numbered floor 1, you determine

which floor you're on. What floor is that? (*b*) Before you make the phone call, however, you realize you've forgotten that the sound of the impact at the bottom travels up the shaft at 340 m/s, and so you redo your calculation. What floor number do you tell the rescuers? ●●

89. Starting from rest, a motorcycle moves with constant acceleration along a highway entrance ramp 0.22 km long. The motorcycle enters the traffic stream on the highway 14.0 s later at 31.0 m/s. (*a*) What is the magnitude of the acceleration along the ramp? (*b*) How fast is the motorcycle moving 9.5 s after it begins to accelerate? (*c*) How far along the ramp is the motorcycle 9.5 s after it begins to accelerate? ●●

90. Hold a dollar bill with its short sides parallel to the floor, and have a friend hold the thumb and forefinger of one hand on either side of the bottom edge, with the thumb and finger about 20 mm apart. Offer your friend the following deal: He can keep the dollar if he can catch it between his thumb and finger after you release it. Most people fail to catch the bill. Use this information (and your measurement of the length of the long side of a dollar bill) to calculate a typical minimum reaction time. ●●

91. A bullet is fired directly upward and reaches a height of 45 m. (*a*) What is the muzzle speed of the gun? (*b*) How much time does it take the bullet to reach its maximum height? ●●

92. A rock dropped from the top of a building travels 30 m in the last second before it hits the ground. How high is the building? ●●

93. A ball launched directly upward from the ground at an initial speed of 24.5 m/s hits the ground 5.0 s later. What are (*a*) the ball's upward position 1.0, 2.0, 3.0, and 4.0 s after launch; (*b*) its upward velocity 1.0, 2.0, 3.0 and 4.0 s after launch; (*c*) its average velocity while it is in the air; and (*d*) its average speed for the entire trip? ●●

94. After landing on Mars, you drop a marker from the door of your landing module and observe that it takes 2.1 s to fall to the ground. When you dropped the marker from the module door on Earth, it took 1.3 s to hit the ground. What is the magnitude of the acceleration due to gravity near the surface of Mars? ●●

95. A very elastic rubber ball dropped from a height of 2.1 m rebounds to 88% of its original height. If the ball is in contact with the floor for 0.013 s, what is its average acceleration during that interval? ●●

96. A scout wants to determine the depth of a dark well. She drops a pebble and hears a splash 5 seconds later. If the speed of sound in air is 343 m/s, (*a*) how long does it take the pebble to reach the water? (*b*) How deep is the well? ●●

97. An elevator cable breaks, and the elevator car falls four floors (from rest) before the emergency brakes kick in. If the car falls one more floor before it stops, what is its average acceleration during braking? ●●

98. A car initially traveling with velocity $\vec{v}_i$ slows down at constant acceleration and stops. (*a*) Sketch a graph of $v_x(x)$, the car's x component of velocity as a function of *position*. (*b*) If the car travels a distance ℓ during this braking, after what displacement does it have half its original speed? ●●

99. A theorist has predicted a form of matter, called X-matter, that is gravitationally attracted to other X-matter but gravitationally *repels* ordinary matter. On a planet made of X-matter, an object made of ordinary matter would feel an "antigravitational" acceleration, pointing upward, away from the planet's surface. Such an object released near the ground would move not in free-fall motion but rather in "free-rise motion." Suppose that on this planet, a projectile launched downward from a cliff at an initial speed of 20 m/s is observed 10 s later moving upward at 40 m/s. (*a*) What is the free-rise acceleration on this planet? (*b*) At what instant after launch is the projectile's speed zero? ●●

100. A rocket is fired vertically up at an average acceleration of 2 m/s^2. Taking $t = 0$ at lift-off, (*a*) what is the velocity of the rocket at $t = 1$ min? (*b*) What is its velocity at $t = 2$ min? (*c*) What is the altitude after 30 seconds? ●●

101. From a promontory overhanging a lake, you throw a stone that enters the water vertically at 15.0 m/s. You hear the splash 2.68 s after you release the stone. Given that the speed of sound in air is 340 m/s: (*a*) What was the stone's initial speed? (*b*) Did you throw the stone down or up? (*c*) From what height above the lake surface did you release the stone? ●●

102. Your state's *Driver Handbook* says, "When the vehicle ahead of you passes a certain point such as a sign, count 'one-thousand-one, one-thousand-two, one-thousand-three.' This takes about three seconds. If you pass the same point before you finish counting, you are following too closely." In your car, you are following a truck on the highway, both of you moving at 100 km/h. You are keeping a 3.0-s gap between your front bumper and the rear of the truck. Suddenly a large box appears on the road beneath the rear of the truck. Apparently the truck's axles were tall enough to allow it to pass right over the object, but yours are not! It takes you 0.75 s to react and hit your brakes. (*a*) What is the minimum constant acceleration required if you are not to hit the box? (*b*) What would your answer be if you were keeping a 2.0-s gap between your car and the truck, as is sometimes recommended as safe, under optimal driving conditions, for alert, experienced drivers with good reflexes? Compare and comment on your answers in parts *a* and *b*. ●●

103. A falling rock hits the ground at a speed of 16 m/s. How many seconds were needed for the rock to fall the last 12 m? ●●

104. You want to determine whether a tall tree will fit onto your 75-m-long logging trailer after you cut it down. You throw a rock straight up with a speed great enough that the rock just reaches the top of the tree, releasing the rock 2.0 m above the ground. You observe that the time interval from the instant you release the rock to the moment it hits the ground is 5.0 s. How tall is the tree? ●●● CR

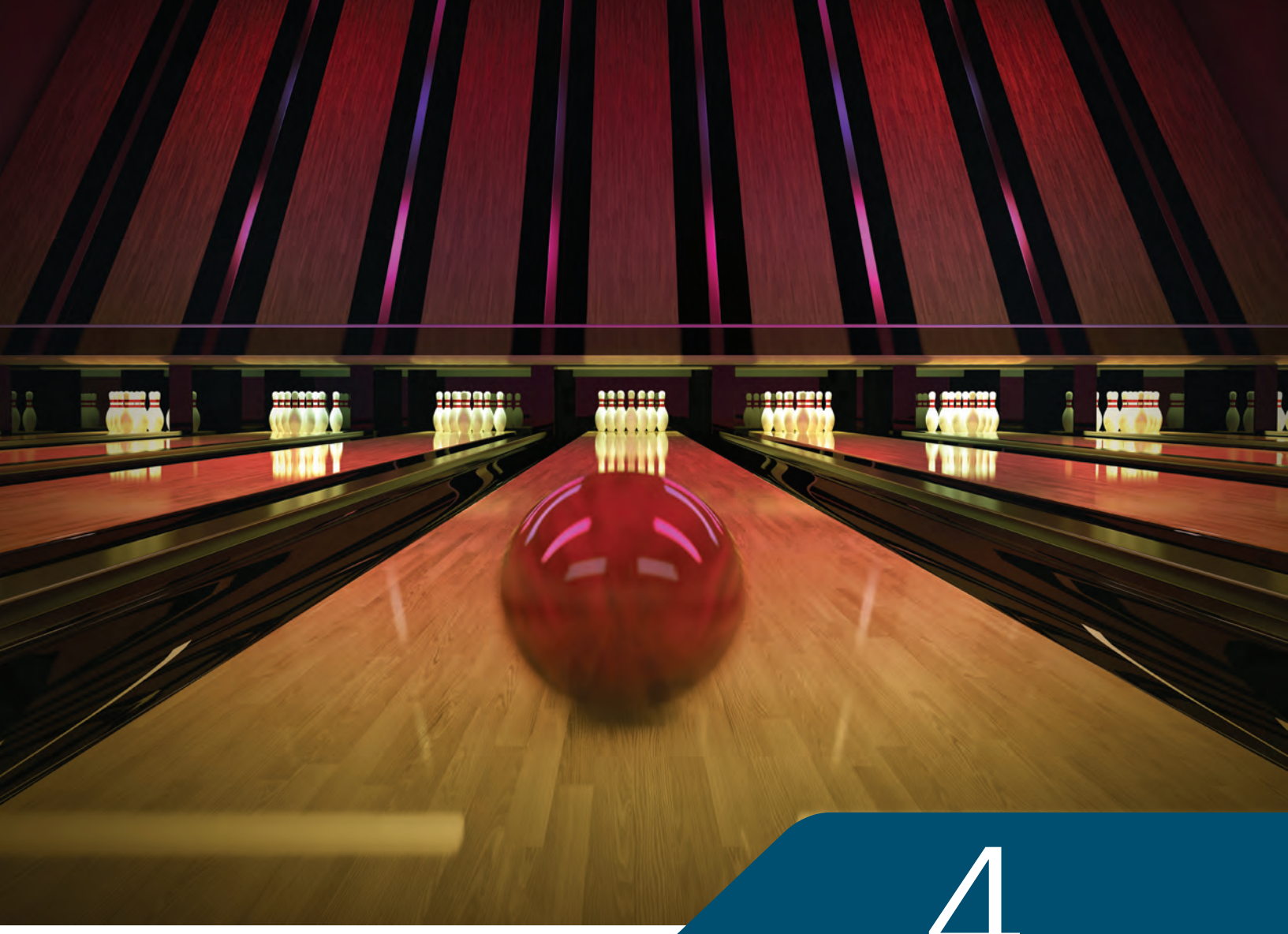
105. Your cousin is always bragging about how high his model rockets go, so you go shopping. None of the brands specify altitude, but you do find three models in your price range with comparable technical information. "Cloudscraper" promises a 3.3-s fuel burn time and 2.9*g* vertical acceleration. "Stratosphere" offers 3.0 s at 3.2*g*, and "Astronaut" claims 2.7 s at 3.6*g*. Which will go highest? ●●● CR

106. You are enjoying a coffee break with your cousin in her highway patrol car when a sports car speeds by at 135 km/h. Your cousin tells you to hold her coffee as she starts the car, and she offers to bet you ten dollars that she can catch the speeder before he reaches the state line. You know that the state line is 2.0 km away along the straight and level highway, and you have read that police cruisers can reach a top speed of 210 km/h within 15 s from a resting start. As the car begins to move, you realize that 5.0 s has already elapsed and that you must decide soon. Should you take the bet? ●●● CR

107. At the drag races, three physics students note that the winning dragster in one category has a speed of 215 mi/h at the end of its quarter-mile run. The run takes 6.30 s, and the car, initially at rest, starts exactly when the timing clock starts. Using these data, and with Chapter 3 as their guide, the students calculate the magnitude of the dragster's acceleration. Although their physics instructor verifies that they have made no numerical errors, the students get three different values for the acceleration. (*a*) What are those values? (*b*) One of the students notices an interesting connection among the three values they obtained. What is that connection? (*c*) Is there a reasonable explanation for the discrepancies found in part *a*? ●●●

108. You are touring a distant planet on which the magnitude of the free-fall acceleration is 65.0% of what it is on Earth. For a little excitement, you jump off a precipice 500 m above the planet's surface. After 5.00 s of free fall, you ignite the jet-pack on your back, changing your acceleration to some new, constant value for the remainder of the fall. You reach the ground 26.0 s after igniting the jet-pack. At what speed do you hit the ground? ●●●

109. A classmate leaves a message on your voice mail betting that you cannot throw a stone high enough so it lands on the roof of a 20-m-high building. As you stare out of your window pondering whether to accept the challenge, the well in the courtyard suddenly gives you an idea. You drop a stone into the well and note that you hear a splash 4.0 s later. You repeat the experiment with another stone, but this time you throw the stone down as fast as you can. This time the splash comes 3.0 s after the stone leaves your hand. You carry out a quick calculation and then you call back your friend. Should you take the bet? ●●● CR



4

Momentum

CONCEPTS

4.1 Friction
4.2 Inertia
4.3 What determines inertia?
4.4 Systems

QUANTITATIVE TOOLS

4.5 Inertial standard
4.6 Momentum
4.7 Isolated systems
4.8 Conservation of momentum

In the preceding two chapters, we developed a mathematical framework for describing motion along a straight line. In this chapter, we continue our study of motion by investigating inertia, a property of objects that affects their motion. The experiments we carry out in studying inertia lead us to discover one of the most fundamental laws in physics—conservation of momentum.

4.1 Friction

Picture a block of wood sitting motionless on a smooth wooden surface. If you give the block a shove, it slides some distance but eventually comes to rest. Depending on the smoothness of the block and the smoothness of the wooden surface, this stopping may happen sooner or it may happen later. If the two surfaces in contact are very smooth and slippery, the block slides for a longer time interval than if the surfaces are rough or sticky. This you know from everyday experience: A hockey puck slides easily on ice but not on a rough road.

Figure 4.1 shows how the velocity of a wooden block decreases on three different surfaces. The slowing down is due to *friction*—the resistance to motion that one surface or object encounters when moving over another. Notice that, during the interval covered by the velocity-versus-time graph, the velocity decrease as the block slides over ice is hardly observable. The block slides easily over ice because there is very little friction between the two surfaces. The effect of friction is to bring two objects to rest with respect to each other—in this case the wooden block and the surface it is sliding on. The less friction there is, the longer it takes for the block to come to rest.

Figure 4.1 Velocity-versus-time graph for a wooden block sliding on three different surfaces. The rougher the surface, the more quickly the velocity decreases.

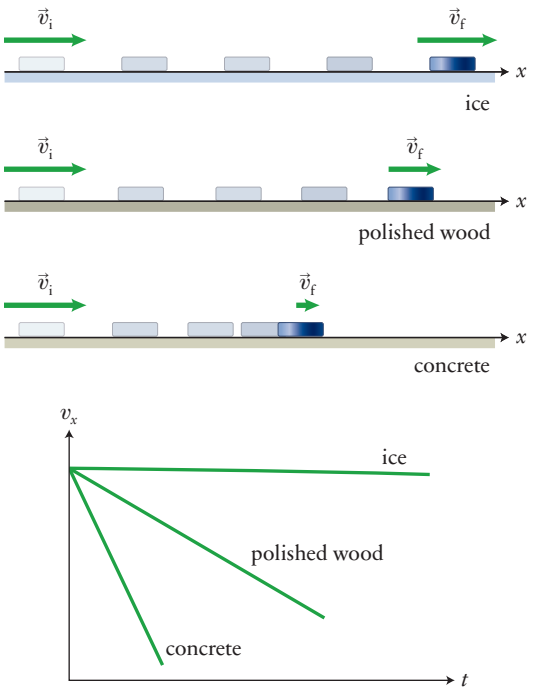


Figure 4.2 Low-friction track and carts used in the experiments described in this chapter.



You may wonder whether it is possible to make surfaces that have no friction at all, such that an object, once given a shove, continues to glide forever. There is no totally frictionless surface over which objects slide forever, but there are ways to minimize friction. You can, for instance, float an object on a cushion of air. This is most easily accomplished with a low-friction track—a track whose surface is dotted with little holes through which pressurized air blows. The air serves as a cushion on which a conveniently shaped object can float, with friction between the object and the track all but eliminated. Alternatively, one can use wheeled carts with low-friction bearings on an ordinary track. **Figure 4.2** shows low-friction carts you may have encountered in your lab or class. Although there is still some friction both for low-friction tracks and for the track shown in Figure 4.2, this friction is so small that it can be ignored during an experiment. For example, if the track in Figure 4.2 is horizontal, carts move along its length without slowing down appreciably. In other words:

In the absence of friction, objects moving along a horizontal track keep moving without slowing down.

Another advantage of using such carts is that the track constrains the motion to being along a straight line. We can then use a high-speed camera to record the cart's position at various instants, and from that information determine its speed and acceleration.

CONCEPTS

4.1 (*a*) Are the accelerations of the motions shown in Figure 4.1 constant? (*b*) For which surface is the acceleration largest in magnitude?

4.2 Inertia

We can discover one of the most fundamental principles of physics by studying how the velocities of two low-friction carts change when the carts collide. Let's first see what happens with two identical carts. We call these *standard carts* because we'll use them as a standard against which to compare the motion of other carts. First we put one standard cart on the low-friction track and make sure it doesn't move. Next we place the second cart on the track some distance from the first one and give the second cart a shove toward the first. The two carts collide, and the collision alters the velocities of both.

Figure 4.3 shows a high-speed video sequence of such a collision. Cart 1 is initially at rest, and consequently its position doesn't change in the first seven frames. Cart 2 approaches cart 1 from the left and collides with it 60 ms after the beginning of the video sequence. The collision causes cart 1 to move to the right and cart 2 to stop dead in its tracks (or, rather, track).

Figure 4.3 (*a*) Two identical carts collide on a low-friction track. The length of track visible in each frame is 0.40 m. Cart 1 is initially at rest; cart 2 collides with it in the middle of the sequence. (*b*) Two curves, one marking the position of the rear of cart 1 and the other marking the position of the front of cart 2, superimposed on the video frames.

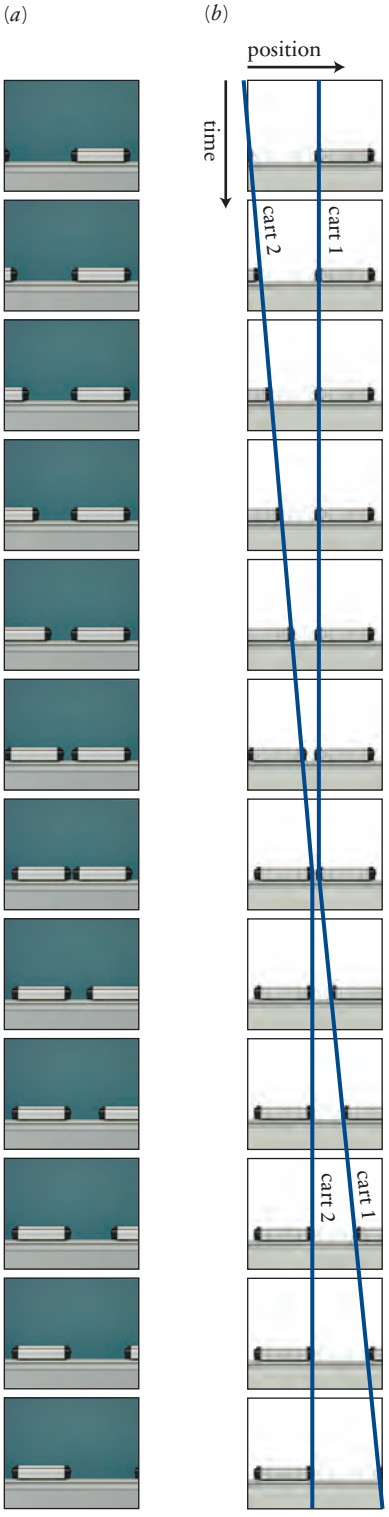
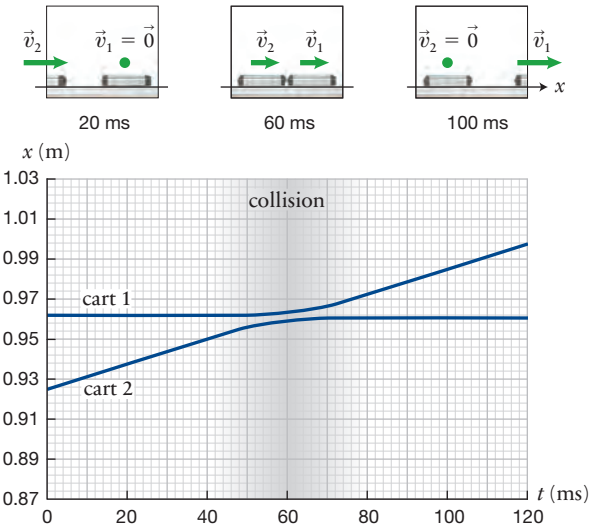


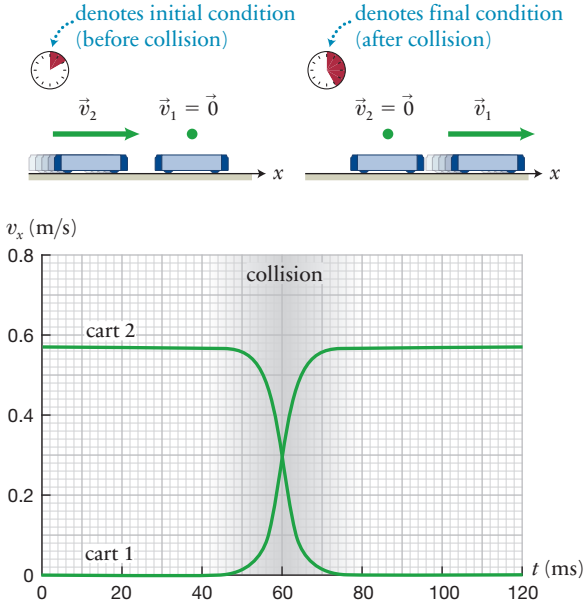
Figure 4.4 Position-versus-time graph for the two carts in Figure 4.3. The curves correspond to the curves of Figure 4.3*b*. All 120 original frames in the sequence were used to obtain this graph.



Measuring the positions of the two carts at various instants gives us the $x(t)$ curves shown in **Figure 4.4**, where the shaded region shows the time interval during which the collision took place. Although the collision appears "instantaneous" to any observer, it takes about 10 ms for the motion of the carts to adjust.

Figure 4.5, which shows the velocities of the two carts, tells us that initially the velocities are constant: 0 for

Figure 4.5 Velocity-versus-time graph for the carts of Figure 4.3 before and after the collision. Before the collision, cart 1 is at rest; after the collision, cart 2 is at rest and cart 1 moves at the initial velocity of cart 2.



CONCEPTS

Figure 4.6 Velocity-versus-time graph for two identical carts before and after a collision on a low-friction track. Both carts are initially moving in the same direction.

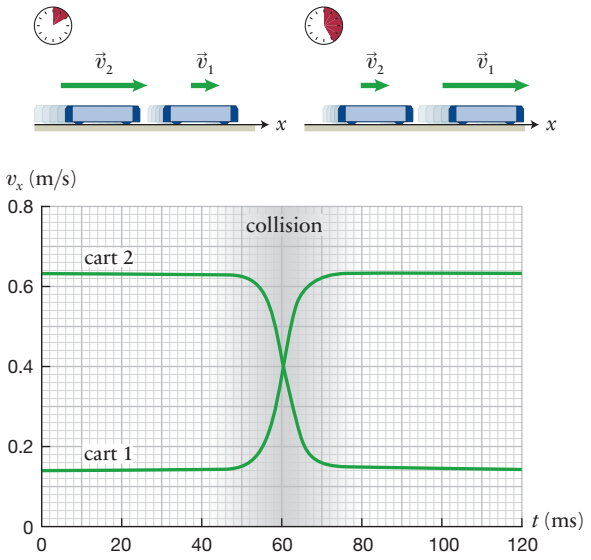
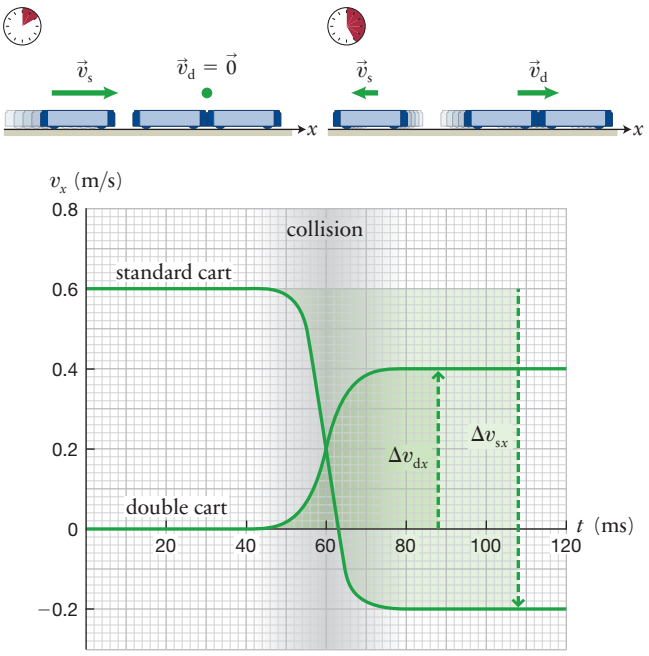


Figure 4.7 Velocity-versus-time graph for a standard cart and a double cart before and after the two collide on a low-friction track. The collision sets the double cart into motion and reverses the direction of travel of the standard cart.



cart 1 and about +0.58 m/s (the plus sign indicating motion to the right) for cart 2. After the collision, the velocities are interchanged: Cart 1 now moves to the right at +0.58 m/s, and cart 2 has come to a complete stop. We can also see this interchanging of velocities in Figures 4.3 and 4.4, where the slopes of the two $x(t)$ curves are interchanged after the collision.

We can repeat the experiment, this time giving cart 2 a harder shove, and then repeat it again, this third time giving cart 2 just a gentle nudge. What we discover is that, *no matter what the initial velocity of cart 2 is*, the collision always interchanges the two velocities.

Further experiments show that it is not necessary that cart 1 be initially at rest. **Figure 4.6** shows what happens when both carts are moving in the same direction at the instant of collision: Once again the collision interchanges the two velocities.

4.2 What is the change in velocity of (*a*) cart 1 and (*b*) cart 2 in Figure 4.6? (*c*) What do you notice about your two answers?

You can repeat this experiment with many different initial velocities, with the carts moving in the same direction or in opposite directions, and you will always observe that the collision interchanges the velocities of the carts. Moreover, Checkpoint 4.2 shows that, if the velocity of one cart increases by a certain amount as a result of the collision, then the velocity of the other cart decreases by exactly the same amount.

To determine whether the volume of material that makes up each cart affects the motion, let's fasten two standard carts together so that the size of this unit is twice the size of the other cart we are going to use, which we continue to call the standard cart. With this double cart at rest on the track, we shove the standard cart toward it at a speed of 0.60 m/s. This time the moving (standard) cart does not come to a complete stop; instead, the collision reverses its direction of travel, so that after the collision it moves to the left with a speed of 0.2 m/s, as **Figure 4.7** shows. After the collision, the initially stationary double cart, being more difficult to set in motion than the single standard cart, moves to the right, as before, but now its speed is lower than in the earlier experiments (0.40 m/s versus about 0.60 m/s). This result makes sense because we already know from experience that more massive objects are harder to set into motion than less massive ones (other than by the force of gravity)—it's easier to throw a small stone than a large boulder.

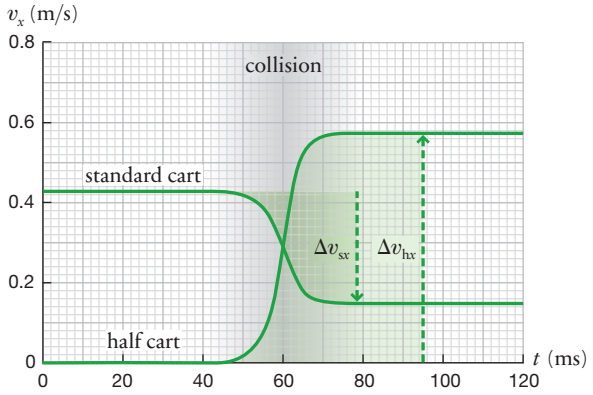
The x component of the velocity of the double cart changed by $\Delta v_{dx} = +0.40\ \text{m/s} - 0 = +0.40\ \text{m/s}$, and the x component of the velocity of the standard cart changed by $\Delta v_{sx} = -0.20\ \text{m/s} - (+0.60\ \text{m/s}) = -0.80\ \text{m/s}$.* The magnitude of the double cart's velocity change, $|\Delta v_{dx}| = 0.40\ \text{m/s}$, is half that of the standard cart, $|\Delta v_{sx}| = 0.80\ \text{m/s}$.

We can repeat this experiment and vary the initial speeds and directions of motion, but we would continue

*Because we must keep track of the changes in velocity of two colliding objects, we need an additional subscript. For example, $\Delta v_{dx} = v_{dx,f} - v_{dx,i}$ is the change in velocity for the double cart. The final subscript, separated by a comma, designates the instant: f for final and i for initial. The middle subscript indicates that we are dealing with the x component. The first subscript designates the object (d for double cart).

CONCEPTS

Figure 4.8 Velocity-versus-time graph for a standard cart and a half cart before and after the two collide on a low-friction track.



to observe that *no matter how the carts move (or do not move) initially, the magnitude of the velocity change of the double cart is always half that of the standard cart.* Also, the two velocity changes are always in opposite directions, which means the x component of the velocity of one cart increases and that of the other cart decreases.

Now let us repeat the collision once again, this time with the original cart 1 cut in half so that it is half the size of the other cart we'll be using. With the half cart at rest on the track, we push a standard cart toward it. As **Figure 4.8** shows, the magnitude of the velocity change of the half cart is twice the magnitude of the velocity change of the standard cart: $|\Delta v_{hx}| = 0.58$ m/s and $|\Delta v_{sx}| = 0.29$ m/s.

A pattern is developing here: For a group of objects all made of the same material, the motion of larger objects is harder to change than the motion of smaller objects. The larger objects put up more resistance when we try to change their velocity. This tendency of an object to resist a change in its velocity is called **inertia.**

Inertia is a measure of an object's tendency to resist any change in its velocity.

The results of our experiments are summarized in **Table 4.1**. As you can see, the *ratio of the inertias* of the two carts is equal to the inverse of the ratio of their velocity changes. In experiment 2, the double cart inertia is twice the standard cart inertia, and the ratio $|\Delta v_{dx}| : |\Delta v_{sx}|$ is 0.5. In experiment 3, the half cart inertia is half that of the standard cart, and the ratio $|\Delta v_{hx}| : |\Delta v_{sx}|$ is 2.

The data in Table 4.1 suggest that we could now take a new cart of unknown inertia, let a standard cart collide with it, and then *determine* from the changes in velocity the inertia of the new cart relative to that of the standard cart.

Table 4.1 Ratio of Velocity Changes in Collisions Between Two Carts

Experiment	Cart 1	Cart 2	$\lvert\Delta v_{1x}\rvert : \lvert\Delta v_{2x}\rvert$
1	standard	standard	1
2	double	standard	0.5
3	half	standard	2

An oil supertanker has an enormous inertia. Even at full throttle it can take as long as 10 min to bring a large ship to a halt, which explains why ships like the *Titanic* and the *Exxon Valdez* could not avoid a collision in spite of seeing the danger well ahead of time.

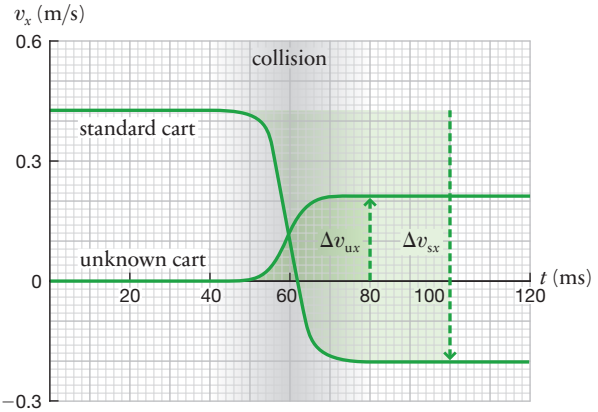


4.3 The x component of the final velocity of the standard cart in Figure 4.8 is positive. Can you make it negative by adjusting this cart's initial speed while still keeping the half cart initially at rest?

Figure 4.9 is a graph of the collision between some unknown cart and a standard cart—both solid and made of the same material. The magnitude of the unknown cart's change in velocity is one-third that of the standard cart (Checkpoint 4.4). We conclude from this result that the unknown cart's inertia is three times that of the standard cart. In other words, the amount of material in the unknown cart is three times that in the standard cart.

4.4 Verify that $|\Delta v_{ux}|/|\Delta v_{sx}| \approx 1/3$ for the two carts in Figure 4.9.

Figure 4.9 Velocity-versus-time graph for a standard cart and a cart of unknown inertia before and after the two collide on a low-friction track. The unknown inertia can be determined from the ratio of the carts' velocity changes.



CONCEPTS

4.3 What determines inertia?

Having established that an object's inertia (its resistance to a change in velocity) is proportional to the amount (roughly, volume) of material in the object, we now come to the next important question: Is it only the amount of material that determines inertia? So far, all the carts we've been studying have been made of the same material. Would our results in experiment 1 be the same if we replaced one of the standard carts with a cart of the same volume of material but made of a different material—plastic instead of metal, say? As you probably know from experience, the answer is *no:* For example, a lead ball is more difficult to set in motion than a rubber ball of equal volume.

So let's make two carts of identical amounts of material, one of plastic and the other of metal, and see what happens when the two collide. The result, shown in **Figure 4.10**, is very different from the result for two identical carts (Figure 4.5).

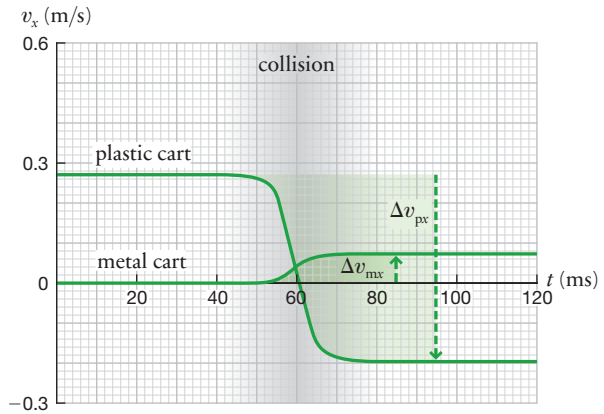
4.5 What is the ratio of the x components of the change in velocity for the plastic and metal carts, $\Delta v_{px}/\Delta v_{mx}$, in Figure 4.10?

As we expect, the x component of the velocity of the plastic cart changes by a much larger amount than that of the metal cart (but it is still opposite in sign). In other words, it is easier to change the motion of the plastic cart than the motion of the metal cart. Given that inertia is a measure of an object's tendency to resist any change in its velocity, we conclude that an object's inertia is determined not just by the volume of material it contains but also by what that material is.

What about other physical properties—shape, say, or the smoothness of an object's surface? Experiments show that inertia does not depend on any of these other properties:

The inertia of an object is determined entirely by the type of material of which the object is made and by the amount of that material contained in the object.

Figure 4.10 Velocity-versus-time graph for two colliding carts that have the same volume but are made of different materials.



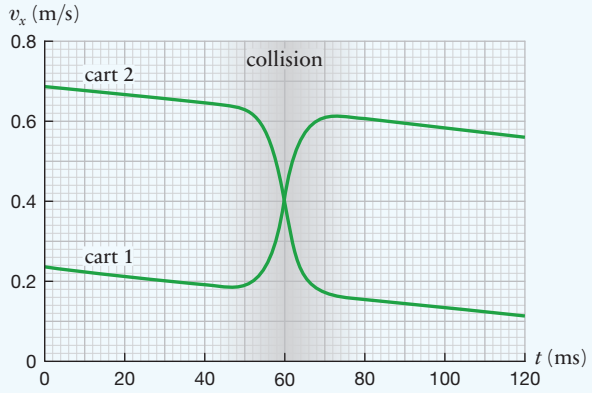
Once we have chosen one object as the standard for inertia, we can determine the inertia of any other object—whatever its size, shape, or composition—by letting it collide with our standard and measuring the ratio of the changes in the velocity of the two objects.

4.6 Is the inertia of the cart of unknown inertia in Figure 4.9 greater or less than that of the standard cart?

EXAMPLE 4.1 Friction and collisions

Figure 4.11 shows the $v_x(t)$ curves for a collision between two identical carts moving not on a low-friction track but rather on a rough surface, so that friction affects their motion. Are the changes in the velocity of the carts caused by the collision still equal in magnitude?

Figure 4.11 Example 4.1.



1 GETTING STARTED Looking at the graph, I'm tempted to say yes because if I tilt Figure 4.11 a bit, it looks very similar to Figure 4.6. Determining the change in velocity due to the collision, however, is not as straightforward because friction also affects the velocities. So, because the velocities of the carts before the collision decrease due to friction, it is not possible to read off a well-defined "initial" value of the velocity. The shape of the graph suggests that friction continues to decrease the velocities in the same way during and after the collision, so I have to determine a way to separate the change in velocity due to friction from the change due to the collision.

2 DEVISE PLAN Because the effect of friction is a steady decrease in velocity, I can extrapolate the $v_x(t)$ curve for cart 1 to calculate what its velocity would have been at $t = 80$ ms if the collision had not taken place. To determine the change in velocity due to the collision, I then read off the actual value of the velocity of cart 1 at $t = 80$ ms and subtract the two velocities. Then I repeat the procedure for cart 2 and compare the changes in velocity.

3 EXECUTE PLAN By placing a ruler along the first part of the $v_x(t)$ curve for cart 1, I observe that the curve lies in line with the second part of the $v_x(t)$ curve for cart 2.

(Continued)

CONCEPTS

Reading off the value of the curve at $t = 80$ ms gives a value of +0.15 m/s. Repeating this procedure for cart 2, I obtain a value of +0.60 m/s. Now I can determine the velocities the carts would have had at $t = 80$ ms if the collision had not taken place. As the graph shows, however, the carts do collide and their velocities at $t = 80$ ms are interchanged. The magnitudes of the changes in velocity caused by the collision are thus 0.45 m/s and 0.45 m/s. These velocity changes are equal in magnitude. ✔

❹ **EVALUATE RESULT** Extrapolation of the velocity curves in Figure 4.11 shows that the collision interchanges the velocities of the two carts, just as in a collision between identical carts in the absence of friction. The interchanging means that the changes in velocity must be of equal magnitude, as I found numerically.

As this example shows, the magnitude of the changes in velocity in a collision is the same even in the presence of friction. This means that even in the presence of friction, we can determine an object's inertia from the magnitude of the velocity changes in a collision.

4.4 Systems

In the preceding chapters we were always concerned with just a single object—a puck, a car, and so on. Starting with this chapter, however, most situations we encounter deal with any number of objects that interact with one another. To analyze such situations, we shall usually focus on one or more principal objects—a person climbing a mountain, two carts colliding, gas atoms in a container. The first step in any analysis therefore is to separate the object(s) of interest from the rest of the universe:

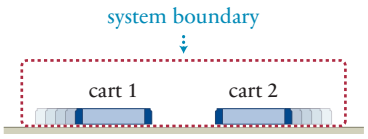
Any object or group of objects that we can separate, in our minds, from the surrounding environment is a system.

For example, when considering a collision between two carts on a low-friction track, we might consider both carts together as the system. When someone throws a ball and we are interested in only the motion of the ball, a logical choice of system might be just the ball by itself. Once we have chosen the ball as our system, everything else—the thrower, the air around the ball, Earth—is outside the system and constitutes its **environment.** This imaginary separation of the universe into two parts—objects within the system and everything in the rest of the universe—makes it possible to develop simple but powerful accounting procedures.

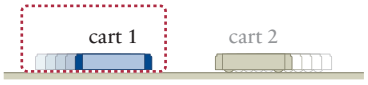
When solving problems, you should imagine a boundary enclosing the objects of interest. What is inside this well-defined boundary constitutes your system and what is outside is the environment. The region of space enclosed in the system boundary may be quite large and contain within it a great many processes or activities, or it may be very small and contain an inert piece of material undergoing no change. The precise shape of the system boundary is unimportant; what counts is that it clearly separates the objects inside the system from those outside. To make this separation explicit, it will generally be helpful to make a pictorial representation of the objects within the system. In your sketch, you should draw a boundary around the objects that constitute your system.

Figure 4.12 Two choices of system for carts colliding on a track.

(*a*) Choice 1: system consists of both carts



(*b*) Choice 2: system consists of one cart



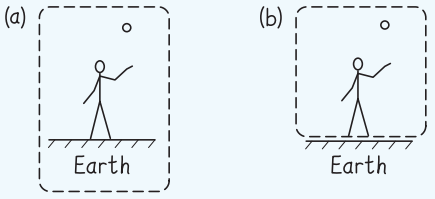
There may be more than one way to separate the system and the environment in any given situation. For two carts colliding on a low-friction track, as illustrated in **Figure 4.12**, we can define our system as containing both carts (in which case the track is part of the environment) or just one cart (in which case the track and the other cart are part of the environment). Deciding what to include in the system will be dictated by the information you wish to learn. Often the choice is obvious; sometimes experience may help you make this decision. More important, once you decide to include a certain object in the system, *it must stay that way throughout the analysis.* Failing to make a consistent choice of system is a frequent source of error.

EXERCISE 4.2 Choosing a system

Cite at least two possible choices of system in each of the following situations. For each choice, make a sketch showing the system boundary and state which objects (excluding air) are inside the system and which are outside. (*a*) After you throw a ball upward, it accelerates downward toward Earth. (*b*) A battery is connected to a light bulb that illuminates a room.

SOLUTION (*a*) The description of the situation mentions three objects: the ball, Earth, and you. One option is to include all three of them in the system (**Figure 4.13*a***). As a second choice, I include you and the ball in the system (Figure 4.13*b*). ✔

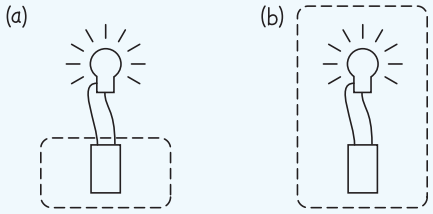
Figure 4.13



CONCEPTS

(*b*) Again I have three objects: the battery, the light bulb, and the room. I can choose just one of them—the battery—as my system (**Figure 4.14*a***) or two of them—the battery and the light bulb (Figure 4.14*b*). ✔

Figure 4.14

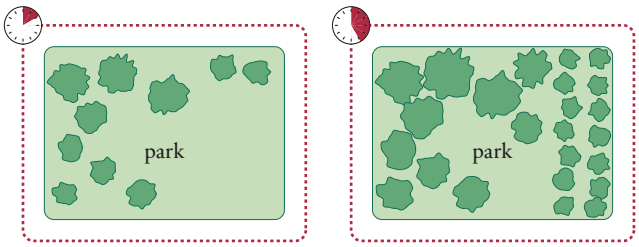


Note that my choices of system are arbitrary. Nothing in the problem prescribes the choice of system. If you tried this problem on your own before looking at my solution and you made different choices, then your answer is just as "correct" as mine!

Defining a system tells us nothing whatsoever about what is happening within it. It is simply a tool to help us set up an accounting scheme. Once we have chosen a system, we can study how certain quantities associated with the system change over time by determining the value of these quantities at the beginning and end of a time interval.

We shall in particular be interested in **extensive quantities**—that is, quantities whose value is proportional to the size or "extent" of the system. More specifically, if we divide the system into a number of pieces, then the sum of an extensive quantity for all the separate pieces is equal to the value of that quantity for the entire system. The number of trees in a park, for example, is extensive: If we divide the park into two parts and add the number of trees in each part, then we obtain the number of trees in the park. The price per gallon of gasoline is not an extensive quantity: If we divide a tankful of gas into two parts and add the price per gallon for the two parts, then we obtain twice the price per gallon for the entire tank. Quantities that do not depend on the extent of the system are **intensive quantities.**

Figure 4.15 System diagram of a park. The number of trees in the park is an extensive quantity.



4.7 Are the following quantities extensive or intensive: (*a*) inertia, (*b*) velocity, (*c*) the product of inertia and velocity?

Only four processes can change the value of an extensive quantity: input, output, creation, and destruction. To see this, consider **Figure 4.15**, which schematically represents a park's initial and final conditions over a certain time interval. The edge of the park is the system boundary, and the park itself—trees and all—is the system. Diagrams that show a system's initial and final conditions are called **system diagrams;** we shall use such diagrams throughout this book. The change in the number of trees over the time interval is given by

$$\text{change} = \text{final tree count} - \text{initial tree count}.$$

The number of trees can change because new trees grow (creation) and old trees die or are taken down (destruction). Alternatively, new trees can be brought into the park (input) and some trees can be moved out of the park (output):

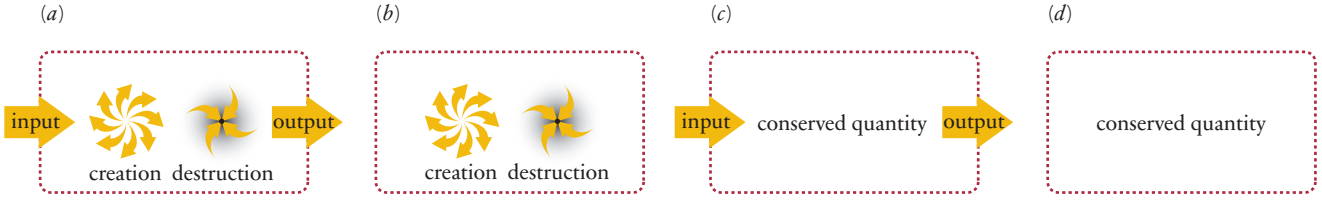
$$\text{change} = \text{input} - \text{output} + \text{creation} - \text{destruction}.$$

This equation is illustrated graphically in **Figure 4.16*a***. If we understand how these four processes affect the trees, we can explain (or predict) any change in their number.

The accounting is often simplified by constraints put on the system. For example, if we build a fence around the park so that no trees can be transported into or out of the park, then the number of trees can change only

CONCEPTS

Figure 4.16 The effect of input, output, creation, and destruction on the extensive quantities of a system.



Extensive quantities can be changed by input, output, creation, and destruction.

If there is no input or output, the quantity can be changed only by creation or destruction.

If an extensive quantity cannot be created or destroyed, it is said to be conserved.

If there is no input or output, a conserved quantity cannot change.

because of creation and destruction. If there is no transfer of the extensive quantity under consideration across the boundary of the system, the change in that property can be due to only creation and destruction (Figure 4.16*b*):

$$\text{change} = \text{creation} - \text{destruction}.$$

Under certain circumstances we can also exclude creation and destruction. Suppose, for example, that we are counting the number of indestructible benches in the park rather than trees. The number of these benches changes only when benches are transported into or out of the park across the park boundary (unless there is a factory making indestructible benches in the park—an unlikely assumption). Any extensive quantity that cannot be created or destroyed is said to be **conserved.** The value of a conserved quantity can change only due to transfer of that quantity across the system boundary (Figure 4.16*c*):

$$\text{change} = \text{input} - \text{output}.$$

For a conserved quantity in a system for which there is no transfer, things are even more simple: Its value cannot change at all (Figure 4.16*d*). Conserved quantities play an important role in physics because their accounting is greatly simplified. In systems for which there is no transfer of a conserved quantity across the system boundary, this quantity does not change *regardless of the processes occurring inside or outside the system* (see the box "Conservation").

EXERCISE 4.3 Accounting principles

(*a*) Classify and give examples of the kinds of processes that can change (*i*) the number of loaves of bread in a bakery, (*ii*) the number of Lego pieces inside a house, and (*iii*) the number of coins in a safe that remains locked. (*b*) For each of these three cases, what is the system? Is the transfer of the quantity of interest across the system boundary possible? Is the quantity of interest conserved?

SOLUTION (*a*) (*i*) Input: a delivery of loaves of bread to the bakery; output: customers leaving the bakery with loaves they purchased; creation: the baking of a batch of loaves; destruction: the eating of loaves by patrons (inside the bakery; destruction outside the bakery falls under "output" because the loaves first leave the bakery). (*ii*) Input: Lego pieces are brought into the house; output: Lego pieces are taken out of the house; creation: assuming there is no manufacturing of Lego pieces in the house, there is no creation; destruction: for all practical purposes, I can consider the Lego pieces to be indestructible and so there is no destruction. (*iii*) Input: if the safe is locked it's impossible to add coins, so the input is zero; output: zero, too; creation: zero; destruction: zero. (Everything is zero, meaning the number of coins can't change inside a locked safe—which is precisely why people put valuables in safes!) ✔

Conservation

Imagine a sheet of paper of width a and length b. The circumference of this sheet is $2a + 2b$, and the surface area of one side is ab. Suppose we cut the sheet lengthwise into two equal parts. Are there any quantities that are unchanged by the cutting?

Circumference is not a conserved quantity: The sum of the circumferences of the two parts is not equal to the circumference of the single sheet. The circumference of each cut piece is $2\left(\frac{1}{2}a\right) + 2b$, and the combined circumference of the two pieces is $2(a + 2b)$, whereas the original circumference was $2(a + b)$.

Surface area is conserved: The surface area of each cut piece is $\frac{1}{2}ab$, and so the sum of the surface areas of the two pieces is ab, exactly as it was before we cut up the sheet.

Trivial? Well, suppose that instead of cutting the sheet in two, we cut it into 148 equal pieces. Can we say anything about the sum of the circumferences? Not without some brainwork! Conservation of surface area, on the other hand, allows us to state—with no calculations whatsoever—that the sum of the surface areas of all the pieces is still ab.

We have found a conserved quantity: surface area! No matter in how many pieces a sheet of paper is cut, the sum of the surface areas of the pieces is equal to the surface area of the original sheet.

If you are a skeptic (and in science you should always be skeptical), you could calculate the surface area of one of the 148 pieces and verify that the sum of the surface areas of all the pieces is indeed ab. To convince yourself that conservation of surface area is a general law, you might also want to repeat this procedure for a few more cases—say, 79 and 237 pieces.

Once you are convinced, however, it is easy to see the value of the conservation law. Suppose that instead of cutting the sheet neatly into equal pieces, someone tears it into ragged pieces, hands them to you, and asks, "What is the combined surface area of all of these pieces?" Even though an actual determination of the surface area would be difficult at best, you would answer, without hesitation, "ab"!

The idea that certain quantities are unaffected by events is appealing, not only because the notion of conservation intuitively makes sense (*some* things must remain unchanged) but also because it provides simple constraints on the behavior of a system.

Is surface area truly conserved? The answer is *no:* you can burn the sheet and make the entire surface area disappear. As we shall see, only a few quantities are truly conserved.

CONCEPTS

(*b*) (*i*) The bakery and its contents make up the system. Loaves of bread can be taken into and out of the bakery, so transfer is possible. The number of loaves of bread is not conserved because loaves can be created and destroyed. (*ii*) The house and its contents are the system. Lego pieces can be taken into and out of it, so transfer is possible. Assuming that the Lego pieces cannot be fabricated or destroyed, I can consider their number to be conserved in this problem. (*iii*) The safe and its contents are the system. The safe remains locked so nothing can be put into it or taken out, and so transfer is not possible. Coins cannot be made or destroyed, so their number is conserved. ✔

Strictly speaking, indestructible benches in a park, Lego pieces, and coins are not truly conserved—all of them are manufactured somewhere (creation) and all of them can be destroyed (destruction). You may be wondering, then, whether *anything* is conserved. The answer is yes: Several quantities we shall encounter are conserved, giving rise to some of the most fundamental principles in physics.

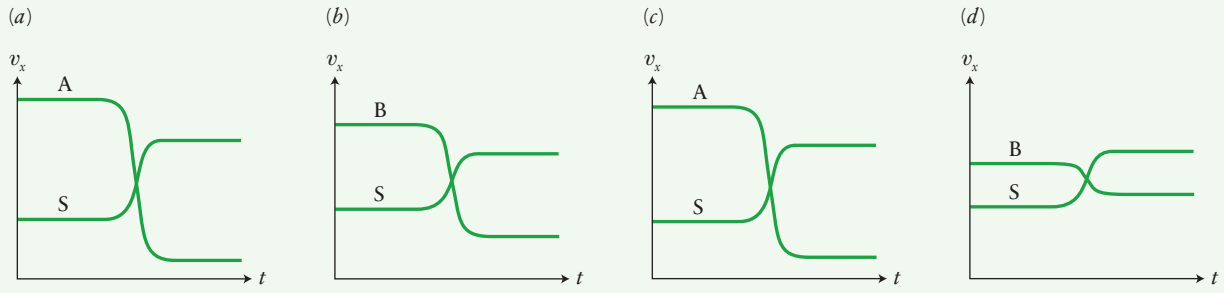


4.8 Which of these quantities is extensive? (*a*) money, (*b*) temperature, (*c*) humidity, (*d*) volume.

Self-quiz

1. Two carts give the same velocity-versus-time graphs when they collide with the same standard cart. Can you conclude from this information that the two carts are identical?
2. The graphs in **Figures 4.17*a*** and 4.17*b* show the effects of carts A and B colliding (separately) with a standard cart S. List the three carts in order of increasing inertia.

Figure 4.17



3. The graphs in Figures 4.17*c* and 4.17*d* show the effects of carts A and B colliding (separately) with a standard cart S. Which has greater inertia, A or B?
4. You do not know whether or not carts A and B of question 3 are made of the same material. Can you tell from the graphs which cart contains a larger quantity of material?
5. Imagine accounting for the number of cattle on a ranch. What do you choose as your system? What are the processes corresponding to input, output, creation, and destruction? Is transfer into and out of the system possible? Is the accountable quantity conserved?

Answers

1. No. Although the identical graphs tell you that the two carts have identical inertias, they need not be identical physically. A large plastic cart, for instance, can have the same inertia as a much smaller metal cart.
2. A and B (identical and less), S (greater). Because the magnitude of the velocity change for S is smaller than that for either A or B, cart S must have greater inertia. The *ratio* of the magnitudes of the velocity changes is about the same in both graphs, however, indicating that A and B have about the same inertias.
3. B. Don't be fooled by the bigger velocity change for S in part *c*; you must first consider each graph separately. Graph *c* tells you that A has less inertia than S because the magnitude of the velocity change for A is *larger* than that for S. Graph *d* tells you that B has greater inertia than S because the magnitude of the velocity change for B is *smaller* than that for S. Hence B must have greater inertia than A.
4. No. Although B has greater inertia, you cannot conclude anything about the quantity of material in it relative to A because you don't know from what materials the two carts are made.
5. The ranch is the system. Input: cattle transported into the ranch; output: cattle transported out of the ranch; creation: cattle born on the ranch; destruction: cattle expiring on the ranch. Transfer is possible because there is input and output, and the number of cattle is not conserved because cattle can be created and destroyed.

4.5 Inertial standard

The inertia of an object is represented by the symbol m (m is for *mass*, a concept that is related to inertia and that we shall examine later on). We shall use the term *inertia* rather than *mass* to underscore the physical meaning of this concept: An object's resistance to a change in its velocity. Inertia is a scalar, which means it does not depend on the orientation of any axis in space. By international agreement, the inertial standard used to be a platinum-iridium cylinder stored at the International Bureau of Weights and Measures in Sèvres, France. The inertia of this standard was defined as one **kilogram,** the basic SI unit of inertia ($m_s \equiv 1\ \text{kg}$).* As of 2019, the kilogram is defined in terms of fundamental constants; to within a precision of 0.003%, it is equal to the inertia of 1 liter of pure water at 4° C.

We can consider an inertial standard to be any "reference" block of metal that is known to have an inertia of 1 kg. Armed with the knowledge gained in Sections 4.1 through 4.3, we can determine the inertia m of any other object. To do so, we make the object we are studying collide with the inertial standard and then measure the changes in velocity. In Section 4.2, we defined the ratio of the inertias of two colliding objects as the inverse of the ratio of the magnitude of their velocity changes, and so we can use this ratio for our inertial standard and our object of unknown inertia:

$$\frac{m_u}{m_s} \equiv -\frac{\Delta v_{sx}}{\Delta v_{ux}} \tag{4.1}$$

or

$$m_u \equiv -\frac{\Delta v_{sx}}{\Delta v_{ux}} m_s. \tag{4.2}$$

Why the minus sign? Well, suppose the object of unknown inertia is a replica of the standard. The left side of Eq. 4.1 then becomes $m_s/m_s = 1$. From the collision in Figure 4.5, we know, however, that the velocity of one of the carts increases (positive Δv_x) while that of the other decreases by the same amount (negative Δv_x). Without the minus sign, Eq. 4.1 would therefore become $1 = -1$. Notice that in all collisions we considered (Figures 4.5 through 4.9), the x components of the velocity change for the two colliding objects have opposite signs. Because this is true for *all* collisions, the minus sign in Eq. 4.1, the defining equation for inertia, means inertia is always a *positive* quantity.

In principle, this simple procedure makes it possible to determine the inertia of any object. Substituting the value $m_s \equiv 1$ kg in Eq. 4.2 allows us to determine the inertia of any other object in kilograms. In practice, it may not be a good idea to let objects collide in order to determine their inertia. (I don't recommend determining the inertia of an expensive glass vase in this way!) In Chapter 13 we shall discuss equivalent but more practical ways to determine the inertia of objects.

EXAMPLE 4.4 Determining the inertia of a stone

A small stone is fastened to the top of a standard cart of inertia 1 kg to form a combination of unknown inertia m_u. A second standard cart is then launched with an initial velocity given by $v_{sx,i} = +0.46$ m/s toward the combination that is initially at rest. After the collision, the x component of the velocity of the cart with the stone is $v_{ux,f} = +0.38$ m/s and that of the standard cart is $v_{sx,f} = -0.08$ m/s. What is the inertia of the stone?

❶ **GETTING STARTED** This collision is much like the one depicted in Figure 4.7 with a standard cart moving toward a cart of greater inertia that is initially at rest. The ratio of the carts' inertias follows from the ratio of the magnitude of their velocity changes.

(Continued)

*You should not associate the unit kilogram with *weight*, a physical property that has to do with gravitation rather than inertia. More about the concept of weight in Chapter 13.

2 DEVISE PLAN To calculate the unknown inertia m_u, I can use Eq. 4.2. I know the x components of the two carts' initial and final velocities, so all the quantities on the right side of the equation are known.

3 EXECUTE PLAN I begin by determining the changes in the x components of the velocity for each cart. For the cart with the stone, $\Delta v_{ux} = v_{ux,f} - v_{ux,i} = +0.38\ \text{m/s} - 0 = +0.38\ \text{m/s}$; for the standard cart $\Delta v_{sx} = v_{sx,f} - v_{sx,i} = -0.08\ \text{m/s} - 0.46\ \text{m/s} = -0.54\ \text{m/s}$. Substituting these values into Eq. 4.2, I get

$$m_u = -\frac{\Delta v_{sx}}{\Delta v_{ux}} m_s = -\frac{-0.54\ \text{m/s}}{+0.38\ \text{m/s}}(1\ \text{kg}) = 1.4\ \text{kg}.$$

This represents the inertia of the stone and the cart as one unit. Given that the inertia of the standard cart is 1 kg, the inertia of the stone must be 0.4 kg. ✔

4 EVALUATE RESULT The inertia of the cart-and-stone combination is greater than 1 kg, as I would expect. If the magnitude of the velocity change of this combination were half that of the standard cart $\left(\frac{1}{2}(0.54\ \text{m/s}) = 0.27\ \text{m/s}\right)$, then the inertia of the cart-and-stone combination would be twice that of a standard cart, and so the inertia of the stone would be 1.0 kg. The magnitude of the velocity change is more than half, so I should expect the stone's inertia to be less than 1.0 kg, which agrees with the answer I found.

4.6 Momentum

Our definition of inertia leads to the definition of another important physical quantity: *momentum*. Let's rewrite Eq. 4.2 in a slightly different form by multiplying both sides by Δv_{ux} and then bringing all terms to the left side:

$$m_u \Delta v_{ux} + m_s \Delta v_{sx} = 0. \tag{4.3}$$

Because the changes in velocity Δv_x can be written in the form $v_{x,f} - v_{x,i}$, we can rewrite this equation as

$$m_u(v_{ux,f} - v_{ux,i}) + m_s(v_{sx,f} - v_{sx,i}) = 0 \tag{4.4}$$

or

$$m_u v_{ux,f} - m_u v_{ux,i} + m_s v_{sx,f} - m_s v_{sx,i} = 0. \tag{4.5}$$

This form of Eq. 4.3, which contains nothing but products of inertias and velocities, is important because it suggests something new. The product of the inertia and the velocity of an object is called the **momentum** (plural: *momenta*) of that object. The conventional symbol for momentum is p, and so we can write

$$\vec{p} \equiv m\vec{v}. \tag{4.6}$$

As you learned in Section 2.6, the product of a vector and a scalar is a vector. Because inertia is a scalar and velocity is a vector, momentum is a vector. The direction of the momentum vector for any object is the same as the direction of the object's velocity vector. The x component of the momentum is the product of the inertia and the x component of the velocity:

$$p_x \equiv mv_x. \tag{4.7}$$

The SI units of momentum are the units of the product mv: $\text{kg} \cdot \text{m/s}$. For example, the magnitude of the momentum of an object of inertia 1 kg moving at a speed of 1 m/s is $1\ \text{kg} \cdot \text{m/s}$. Inertia is an intrinsic property of an object (you can't change it without changing the object), but the value of the momentum of that object, like its velocity, can change.

QUANTITATIVE TOOLS

EXAMPLE 4.5 Bullet and bowling ball

Compare the magnitude of the momenta of a 0.010-kg bullet fired from a rifle at 1300 m/s and a 6.5-kg bowling ball lumbering across the floor at 4.0 m/s.

❶ **GETTING STARTED** Momentum is the product of inertia and velocity. I have to calculate this quantity for both the bullet and the bowling ball and then compare the resulting values.

❷ **DEVISE PLAN** Equation 4.6 gives the momentum of an object. To determine the magnitude of the momentum of an object, I must take the product of the inertia m and the speed v: $p = mv$.

❸ **EXECUTE PLAN** Substituting the values given in the problem statement, I get

$$p_{\text{bullet}} = (0.010\ \text{kg})(1300\ \text{m/s}) = 13\ \text{kg}\cdot\text{m/s} \checkmark$$
$$p_{\text{bowling}} = (6.5\ \text{kg})(4.0\ \text{m/s}) = 26\ \text{kg}\cdot\text{m/s}. \checkmark$$

❹ **EVALUATE RESULT** Surprisingly, the magnitudes of the momenta are very close! I have no way of evaluating momenta because I don't have much experience yet with this quantity. However, the bullet has less inertia and a high speed and the bowling ball has greater inertia and a low speed, so it is not unreasonable that the product of these quantities is similar.

Momentum is a quantitative measure of "matter in motion" and depends on both the amount of matter in motion and how fast that matter is moving. Momentum is very different from inertia. A truck, for example, has greater inertia than a fly (it has a higher resistance to a change in its velocity), but if the truck is at rest and the fly is in motion, then the magnitude of the fly's momentum is larger than that of the truck, which is zero. In Example 4.5, the inertias of the bullet and the bowling ball are very different, yet their momenta are similar. Conceptually you can think of an object's momentum as its capacity to affect the motion of other objects in a collision.

With the definition of momentum, we can rewrite Eq. 4.5 in the form

$$p_{ux,f} - p_{ux,i} + p_{sx,f} - p_{sx,i} = 0. \tag{4.8}$$

If we write $\Delta p_{ux} \equiv p_{ux,f} - p_{ux,i}$ and $\Delta p_{sx} \equiv p_{sx,f} - p_{sx,i}$, Eq. 4.8 takes on the beautifully simple form

$$\Delta p_{ux} + \Delta p_{sx} = 0. \tag{4.9}$$

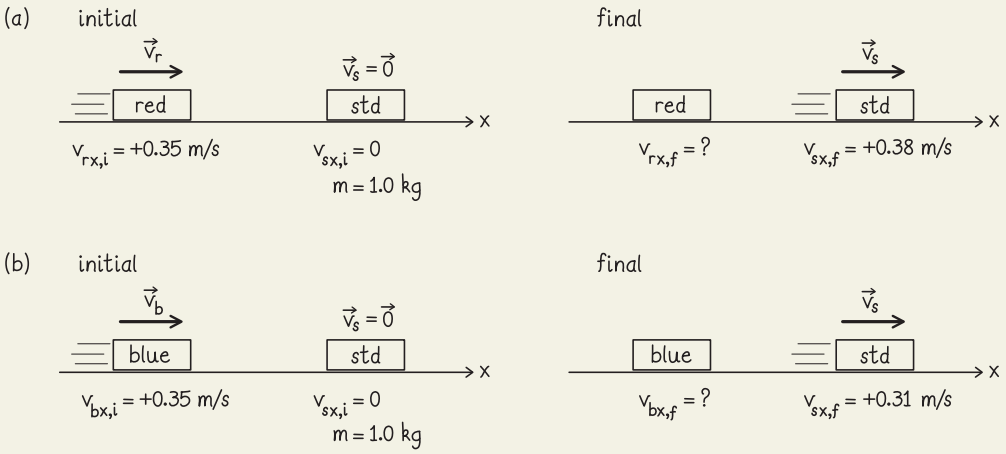
This equation means that, whenever an object of unknown inertia collides with the inertial standard, the changes in the x components of the momenta of the two objects add up to zero. In other words, the change in the x component of the momentum for one object is always the negative of the change for the other.

EXAMPLE 4.6 Collisions and momentum changes

(*a*) A red cart with an initial speed of 0.35 m/s collides with a stationary standard cart ($m_s = 1.0$ kg). After the collision, the standard cart moves away at a speed of 0.38 m/s. What is the momentum change for each cart? (*b*) The experiment is repeated with a blue cart, and now the final speed of the standard cart is 0.31 m/s. What is the momentum change for each cart in this second collision? (*c*) If in the collisions $v_{rx,f} = +0.032$ m/s and $v_{bx,f} = -0.039$ m/s, what are the inertias of the red and the blue carts?

❶ **GETTING STARTED** I begin organizing the information given in the problem in a picture by showing the initial and final conditions for each of the two collisions (**Figure 4.18**).

Figure 4.18



(Continued)

2 DEVISE PLAN In each collision, I know the inertia and the x components of the initial and final velocities of the standard cart, so I can readily determine the change in the x component of its momentum. To determine the change in momentum for the other carts, I can then use Eq. 4.9, which tells me that for each collision the changes in the x components of the momenta of the two carts must add up to zero. For part c, I can use Eq. 4.7 together with the values for the red and blue carts' initial and final velocities and their changes in momenta calculated in parts a and b to determine the inertias.

3 EXECUTE PLAN (a) Because it initially is at rest, the standard cart's initial momentum is zero. The x component of its final momentum is $p_{sx,f} = (1.0\ \text{kg})(+0.38\ \text{m/s}) = +0.38\ \text{kg}\cdot\text{m/s}$, and so the x component of its change in momentum is

$$\Delta p_{sx} = +0.38\ \text{kg}\cdot\text{m/s} - 0 = +0.38\ \text{kg}\cdot\text{m/s}. \checkmark$$

Because the x components of the momentum changes of the two carts must add up to zero (Eq. 4.9), the corresponding x component of the red cart's momentum change is $\Delta p_{rx} = -0.38\ \text{kg}\cdot\text{m/s}$. ✔

(b) Applying the same procedure for this second collision, I get

$$\Delta p_{sx} = (1.0\ \text{kg})(+0.31\ \text{m/s}) - 0 = +0.31\ \text{kg}\cdot\text{m/s};$$

$$\Delta p_{bx} = -0.31\ \text{kg}\cdot\text{m/s}. \checkmark$$

(c) The change in momentum for the red cart is $\Delta p_{rx} = p_{rx,f} - p_{rx,i}$. Using Eq. 4.7, I can write this change in momentum as

$$\Delta p_{rx} = m_r v_{rx,f} - m_r v_{rx,i} = m_r(v_{rx,f} - v_{rx,i})$$

and so

$$m_r = \frac{\Delta p_{rx}}{v_{rx,f} - v_{rx,i}} = \frac{-0.38\ \text{kg}\cdot\text{m/s}}{+0.032\ \text{m/s} - 0.35\ \text{m/s}}$$
$$= 1.2\ \text{kg}. \checkmark$$

Likewise, for the blue cart:

$$m_b = \frac{\Delta p_{bx}}{v_{bx,f} - v_{bx,i}} = \frac{-0.31\ \text{kg}\cdot\text{m/s}}{-0.039\ \text{m/s} - 0.35\ \text{m/s}}$$
$$= 0.80\ \text{kg} \checkmark$$

4 EVALUATE RESULT If either the red or the blue cart had an inertia of 1.0 kg, then the cart would come to rest in the collision. My work shows that the inertia of the red cart is greater than that of a standard cart. Indeed, after the collision it still moves in the same direction it initially moved, as it should for a cart of greater inertia colliding with one of lesser inertia (see, for example, Figure 4.8). In contrast, the inertia of the blue cart is less than that of the standard cart. After the collision it bounces back and travels in the opposite direction, as I would expect (see Figure 4.7).

Equation 4.9 can be written in vectorial form:

$$\Delta\vec{p}_u + \Delta\vec{p}_s = \vec{0} \tag{4.10}$$

and so, in a collision between an object of unknown inertia and the inertial standard, the changes in the momenta of the two objects add to zero.

4.9 In part a of Example 4.6, what are the directions of the changes in momentum, $\Delta\vec{p}_s$ and $\Delta\vec{p}_r$, of the two carts?

4.7 Isolated systems

As we saw in Checkpoint 4.7, momentum is an extensive property—you can add up the momenta of all the objects in a system to obtain the *momentum of the system*. In particular, the momentum of a system of two moving carts is the sum of the momenta of the two individual carts:

$$\vec{p} \equiv \vec{p}_1 + \vec{p}_2. \tag{4.11}$$

With this definition we can begin to develop an accounting scheme for the momentum of a system. In particular, we shall be interested in determining what causes the momentum of a system to change. Let us begin by reexamining a number of situations we have encountered. In Figure 4.19*a* a puck slides to a stop on a wood floor. In Figure 4.19*b* a cart moves at constant velocity on a low-friction track. In Figure 4.19*c* a standard cart collides with another cart on a low-friction track, and in Figure 4.19*d* someone gives a cart moving at constant velocity on a low-friction track a shove. In some cases the velocity (and therefore the momentum) remains constant; in others it changes.

Figure 4.19 Velocity-versus-time and acceleration-versus-time graphs for four situations. The shading indicates nonzero acceleration.

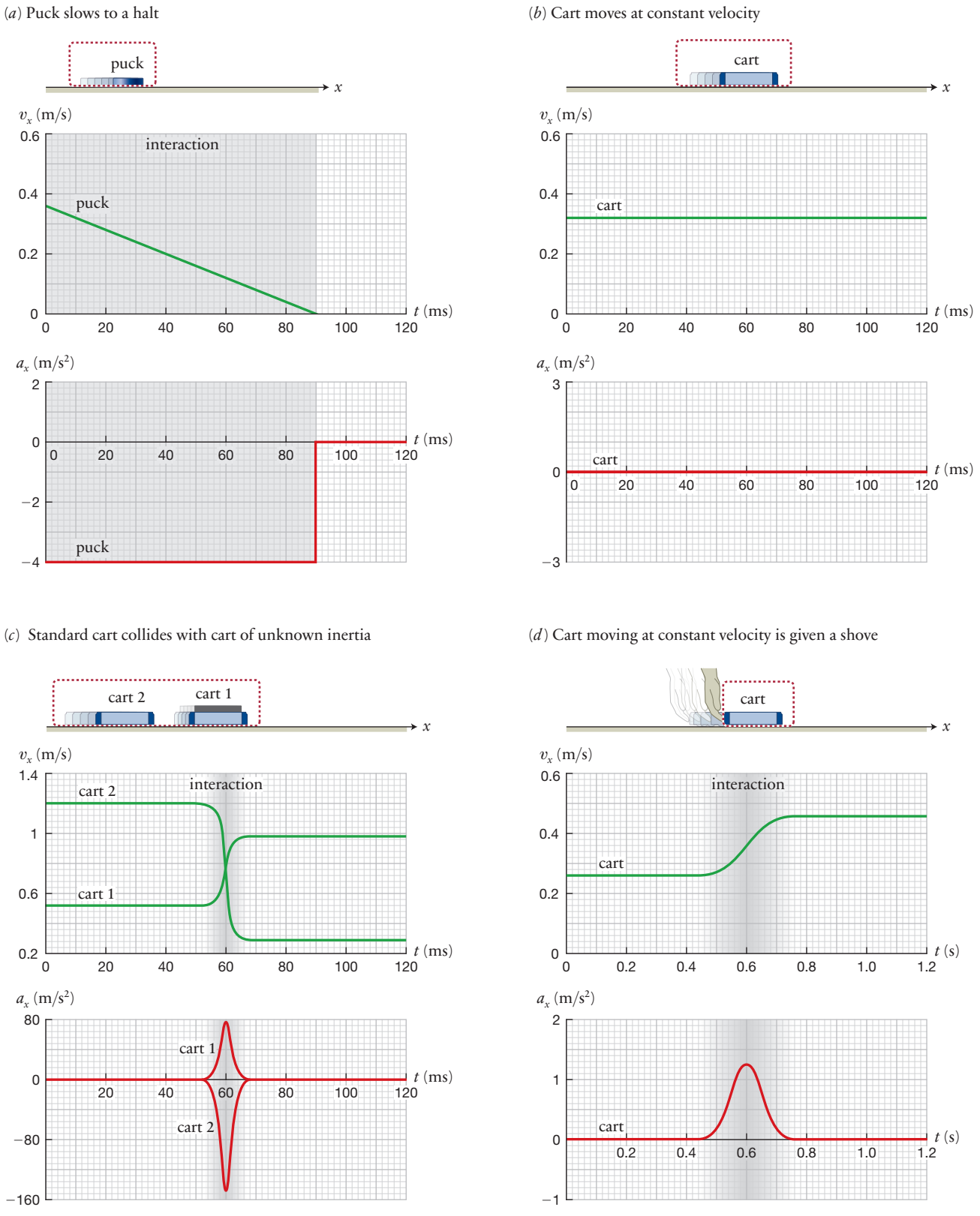


Table 4.2 Interactions and momentum changes in Checkpoint 4.10

Situation	Interacting objects	System	System interacting?	$\Delta\vec{p}$
a	floor ↔ puck	puck	yes	nonzero
b	none	cart	no	$\vec{0}$
c	cart 1 ↔ cart 2	cart 1	yes	nonzero
d	cart 1 ↔ cart 2	carts 1 & 2	no	$\vec{0}$
e	hand ↔ cart	cart	yes	nonzero

4.10 Is the change in the momentum $\Delta\vec{p}$ zero or nonzero for the following choices of system over the 120-ms time interval in Figure 4.19: (*a*) puck in Figure 4.19*a*, (*b*) cart in Figure 4.19*b*, (*c*) cart 1 in Figure 4.19*c*, (*d*) both carts in Figure 4.19*c*, (*e*) cart in Figure 4.19*d*?

In the graphs, I have highlighted the time intervals during which the acceleration of objects within the system is nonzero—that is, the time interval when the velocity and therefore the momentum of objects within the system change. Notice that this time interval always coincides with the time interval during which the object under consideration *interacts* with something else. In Figure 4.19*a* the puck slows down because the puck and the floor rub against (interact with) each other, in Figure 4.19*c* the two carts collide with (interact with) each other, and in Figure 4.19*d* the cart and the hand interact with each other. Conversely, in Figure 4.19*b* the interaction between the track and the cart has been virtually eliminated by reducing friction, and so the acceleration is zero.

In these examples, *interaction* refers to objects touching or rubbing against each other. Interactions are the subject in Chapter 7. For now, we'll use the word *interaction* to mean two objects acting on each other in such a way that one or both are accelerated.

Table 4.2 summarizes the observations from Checkpoint 4.10. Note that whenever the system interacts with something outside the system, the momentum of the system changes. Interactions across the boundary of a system are called *external interactions*. Interactions between two objects inside the system are called *internal interactions*. Looking at Figure 4.19 and Table 4.2, we see that external interactions transfer momentum into or out of the system.

A system for which there are no external interactions is said to be **isolated.** For such a system there is no input or output of momentum, and so the momentum of the system does not change: $\Delta\vec{p} = \vec{0}$. In Figure 4.19*b*, the cart's motion is not influenced by the track, and so the cart constitutes an isolated system. In Figure 4.19*c*, the two carts interact with each other but not with anything outside the system, and so the two carts together constitute a system that is isolated (but either cart by itself does not; see Table 4.3). In Figures 4.19*a* and 4.19*d*, where $\Delta\vec{p} \neq \vec{0}$, the system is not

Table 4.3 Two choices of system for carts colliding

	Choice 1	Choice 2
System:	carts 1 & 2	cart 1
Environment:	track	cart 2 & track
Interactions:	internal	external
System isolated?	yes	no
Momentum changing?	no	yes

QUANTITATIVE TOOLS

isolated. In Figure 4.19*a* the only object in the system (the puck) interacts with the floor, which is outside the system. In Figure 4.19*d*, the only object in the system (the cart) interacts with the hand, which is outside the system.

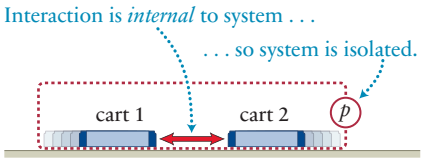
Figure 4.20 and Table 4.3 illustrate two system choices for studying a collision. The system containing both carts is isolated (the interaction between the two carts takes place inside the system; it is an internal interaction), and the momentum of the system is constant. The system containing just cart 1 is not isolated because it interacts with cart 2, which is now part of the environment, and the momentum of the system is not constant.

In isolated systems the analysis of momentum is simplified because we can concentrate on what happens inside the system. The procedure for choosing an isolated system is shown in the Procedure box on the next page and illustrated for a collision between two carts on a low-friction track in **Figure 4.21**. We begin by separating from one another all the objects that could play a role: the carts, the track, and the air. Next we identify all possible interactions: the collision between the two carts, the interaction between the air and the carts, and any possible interaction between the track and the carts. We know that each cart by itself rolls without slowing down on the track, so the interaction between the track and the carts and between the air and the carts does not have any significant effect on the carts' velocity. We therefore eliminate these and retain only the collision between the carts, which does change their velocities.*

Having eliminated all but the interaction between the carts, we can choose an isolated system by enclosing the two carts so that this interaction becomes internal. Finally, we complete the procedure by making a system diagram, showing the initial and final conditions of the isolated system (before and after the collision). In general, the interaction between two colliding objects affects their motion much more strongly than any other interaction, and so we can for all practical purposes always consider the system of two colliding objects to be isolated.

Figure 4.20 The choice of system determines whether the interactions in question are internal or external and hence whether the system is isolated or not. We use the symbol Ⓟ to denote an isolated system, indicating that no momentum crosses the system's boundary.

(*a*) Choice 1: system = both carts



(*b*) Choice 2: system = cart 1

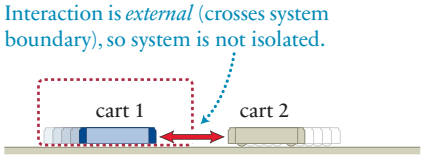
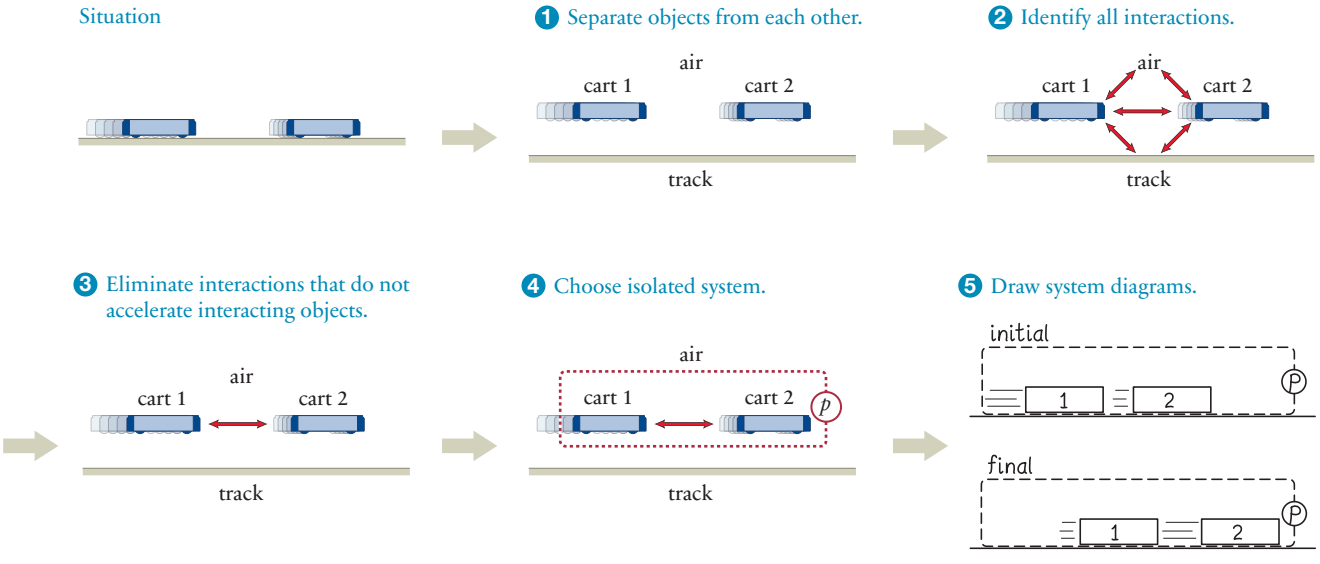


Figure 4.21 Applying the procedure for choosing an isolated system (see the Procedure box) to the collision between two carts on a low-friction track.



*Leaving out the air-cart and track-cart interactions is a simplification, because both *do* have a small effect on the carts' acceleration. However, you can ignore the effects of the these interactions because they are so much smaller than the effects of the collision. Never be afraid to make a simplifying assumption; you can always return to the beginning of a problem and refine your assumptions.

PROCEDURE: Choosing an isolated system

When you analyze momentum changes in a problem, it is convenient to choose a system for which no momentum is transferred into or out of the system (an isolated system). To do so, follow these steps:

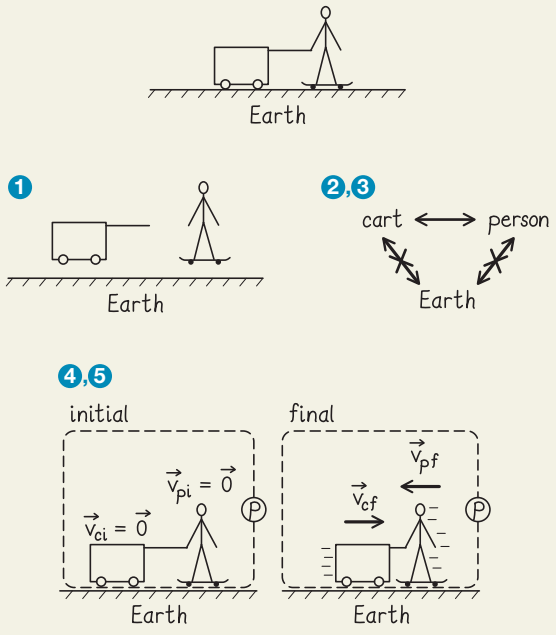
1. Separate all objects named in the problem from one another.
2. Identify all possible interactions among these objects and between these objects and their environment (the air, Earth, etc.).
3. Consider each interaction individually and determine whether it causes the interacting objects to accelerate. Eliminate any interaction that does not affect (or has only a negligible effect on) the objects' accelerations during the time interval of interest.
4. Choose a system that includes the object or objects that are the subject of the problem (for example, a cart whose momentum you are interested in) in such a way that none of the remaining interactions cross the system boundary. Draw a dashed line around the objects in your choice of system to represent the system boundary. None of the remaining interactions should cross this line.
5. Make a system diagram showing the initial and final states of the system and its environment.

EXERCISE 4.7 Who's pulling whom?

A person standing on a skateboard on horizontal ground pulls on a rope fastened to a cart. Both the person and the cart are initially at rest. Use the Procedure box to identify an isolated system and make a system diagram.

SOLUTION See **Figure 4.22**. I begin by separating the objects in the problem: the person, the cart, and Earth. (I could always go into more detail—include the air, the rope, and the skateboard—but it pays to keep things as simple as possible. For that reason, I consider the skateboard to be part of the person and the rope to be part of the cart.) The cart interacts with Earth and the person; the person interacts with the cart and Earth. Ignoring friction in the wheels of the cart, I know that the interaction between the cart and Earth has no effect on any motion, and so I can eliminate it from the analysis. The same holds for the interaction between the person (the skateboard) and Earth. I then draw a boundary around the person and the cart, making the interaction between the two internal. Because there are no external interactions, this system is isolated. Finally I draw a system diagram showing the initial and final conditions of the system with the person and cart initially at rest and then moving. ✔

Figure 4.22



4.11 Imagine sitting on a sled on the slippery surface of a frozen lake. To reposition yourself closer to the back of the sled, you push with your legs against the front end. (*a*) Do you constitute an isolated system? (*b*) How about you and the sled? Ignore the interaction between sled and ice.

4.8 Conservation of momentum

In Section 4.4 we saw that the change in an extensive property can be due to transport into or out of the system and to creation or destruction. In Section 4.7 we saw that whenever we consider a system that is isolated, the momentum of the system is not changing:

$$\Delta\vec{p} = \vec{0} \quad (\text{isolated system}). \tag{4.12}$$

You may be tempted to conclude from this observation that momentum is conserved. All we can conclude, however, is that there is no creation or destruction of momentum *inside the systems we considered*—not that it is impossible to create or destroy momentum under any circumstances.

Because it is a direct consequence of the definition of momentum, Eq. 4.10 must—by definition—hold for *any* object that collides with the inertial standard. So if we take any two objects 1 and 2 and let each collide with the inertial standard on a level, low-friction track, we *always* obtain—no matter how we vary the initial velocities and no matter what the values of inertias m_1 and m_2 are—that

$$\Delta\vec{p}_1 + \Delta\vec{p}_s = \vec{0} \quad (\text{by definition}) \tag{4.13}$$

and

$$\Delta\vec{p}_2 + \Delta\vec{p}_s = \vec{0} \quad (\text{by definition}). \tag{4.14}$$

Suppose now that, instead of making objects 1 and 2 collide with the inertial standard, we make them collide with each other. Should we expect to obtain

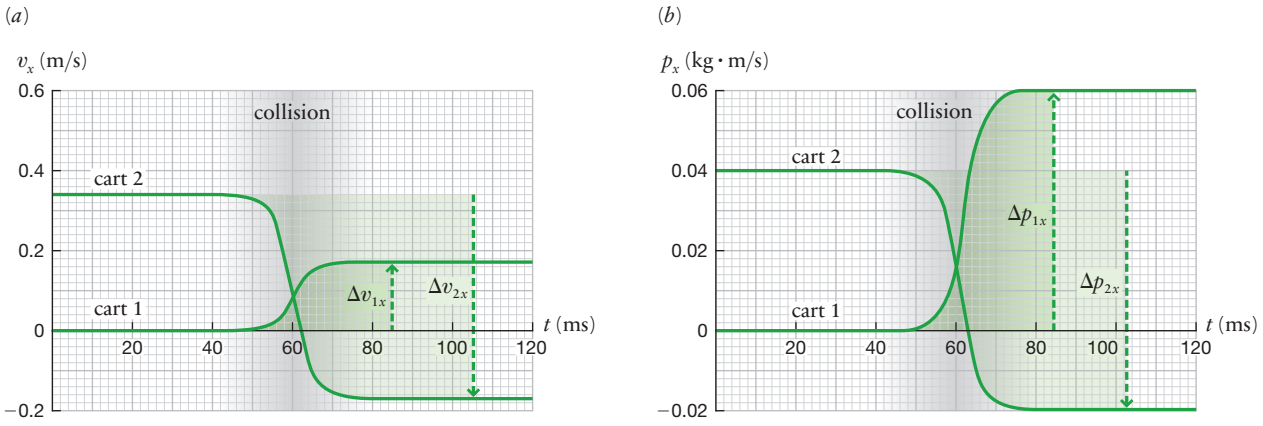
$$\Delta\vec{p}_1 + \Delta\vec{p}_2 = \vec{0}? \tag{4.15}$$

You may be tempted to think that this equation follows from our earlier observations, but that is not the case. The value of $\Delta\vec{p}_s$ in Eqs. 4.13 and 4.14 need not be the same because Δv_{sx}, and therefore $\Delta\vec{p}_s$, vary from collision to collision (compare, for instance, Figures 4.8 and 4.9). So each of these equations holds for a separate collision and therefore we cannot combine them to draw any conclusions about the relationship between $\Delta\vec{p}_1$ and $\Delta\vec{p}_2$. The only way to verify whether Eq. 4.15 holds is to do the measurement.

Figure 4.23 shows the graphs for a collision between two arbitrary carts. To get the momentum values shown in Figure 4.23*b*, I multiplied the x component of the velocity of each cart by the cart's inertia determined by *independently* letting each cart collide with the inertial standard. The change in the x component of the momentum of cart 1 can be obtained by multiplying its inertia ($m_1 = 0.36\ \text{kg}$) with the x components of the initial and final velocities read off from Figure 4.23*a* ($v_{1x,i} = 0$ and $v_{1x,f} = +0.17\ \text{m/s}$):

$$\begin{aligned}\Delta p_{1x} &= m_1(v_{1x,f} - v_{1x,i}) = (0.36\ \text{kg})(+0.17\ \text{m/s} - 0)\\ &= 0.061\ \text{kg}\cdot\text{m/s}.\end{aligned}$$

Figure 4.23 (*a*) Velocity-versus-time graph and (*b*) momentum-versus-time graph of two carts before and after the two collide on a low-friction track. The sum of the momenta is the same before and after the collision. The inertias are $m_1 = 0.36$ kg and $m_2 = 0.12$ kg.



For cart 2 we have $v_{2x,\mathrm{i}} = +0.34$ m/s, $v_{2x,\mathrm{f}} = -0.17$ m/s, and $m_2 = 0.12$ kg, so

$$\begin{aligned}\Delta p_{2x} &= m_2(v_{2x,\mathrm{f}} - v_{2x,\mathrm{i}}) = (0.12\ \mathrm{kg})(-0.17\ \mathrm{m/s} - 0.34\ \mathrm{m/s}) \\ &= -0.061\ \mathrm{kg \cdot m/s}.\end{aligned}$$

Because the two changes in momentum are equal in magnitude and of opposite sign, they add up to zero: $\Delta p_{1x} + \Delta p_{2x} = 0$. Consequently the momentum of the system comprising the two carts does not change:

$$\Delta\vec{p} \equiv \Delta\vec{p}_1 + \Delta\vec{p}_2 = \vec{0}. \tag{4.16}$$

Repeating the experiment with another pair of carts or changing the initial velocities always yields the same result: The momentum of the system comprising any two colliding objects is not changed by the collision. The collision merely transfers some momentum from one cart to the other—one gains a certain amount of momentum, and the other loses the same amount.

4.12 (*a*) How much momentum is transferred in the collisions in Figure 4.23? (*b*) Is this momentum transferred from cart 1 to cart 2, or in the opposite direction?

Experiments show that *any* interaction between two objects—not just collisions—transfers momentum from one object to the other. The sum of the momenta of the interacting objects, however, never changes. No interactions or other phenomena have ever been observed where momentum is created out of nothing or destroyed. We must therefore conclude that momentum is conserved:

Momentum can be transferred from one object to another, but it cannot be created or destroyed.

This statement is one of the most fundamental principles of physics and is often referred to as the **conservation of momentum.**

For an isolated system, conservation of momentum means

$$\Delta\vec{p} = \vec{0} \quad (\text{isolated system}). \tag{4.17}$$

For a system that is not isolated, we have

$$\Delta\vec{p} = \vec{J}, \tag{4.18}$$

QUANTITATIVE TOOLS

where $\vec{J}$ represents the transfer of momentum from the environment to the system. The quantity $\vec{J}$ is called the **impulse** delivered to the system. Like momentum, impulse is a vector and has SI units of kg · m/s. Depending on its direction relative to the momentum, the impulse can increase or decrease the magnitude of the momentum of the system. $\vec{J}$ therefore represents what we called "input" and "output" in Section 4.4.



4.13 (*a*) What is the magnitude of the impulse delivered to cart 1 in Figure 4.23? (*b*) Write the impulse delivered to cart 1 in vector form. (*c*) Does the fact that the change in the momentum of cart 1 is nonzero mean that momentum is not conserved?

Equation 4.18, which we shall use extensively throughout this book, is called the **momentum law.** This equation embodies conservation of momentum because it tells us that the momentum of a system can change only due to a transfer of momentum to the system ($\vec{J}$). If momentum were not conserved, then the right side of this equation would need to contain an additional term to account for changes in momentum due to the creation or destruction of momentum. For an isolated system the impulse is zero, $\vec{J} = \vec{0}$, and the momentum law takes on the form of Eq. 4.17.

EXAMPLE 4.8 Bounce

A 0.20-kg rubber ball is dropped from a height of 1.0 m onto a hard floor and bounces straight up. Assuming the speed with which the ball rebounds from the floor is the same as the speed it has just before hitting the floor, determine the impulse delivered by the floor to the rubber ball.

❶ **GETTING STARTED** I define the ball to be my system in this problem. The impulse delivered to the ball is then given by the change in its momentum (Eq. 4.18). I need to develop a way to determine this change in momentum.

❷ **DEVISE PLAN** To solve this problem I need to first determine the velocity of the ball just before it hits the floor. I therefore break the problem into two parts: the downward fall of the ball and its collision with the floor. I can use Eq. 3.15 to determine the time interval it takes the ball to fall from its initial height (assuming the ball is initially at rest). As it falls, the ball undergoes constant acceleration, so I can use Eq. 3.10 to calculate its velocity just before it hits the floor and, because its speed is not changing as it rebounds, I also know its velocity after it bounces up. Knowing the velocities, I can calculate the ball's momentum before and after the bounce using Eq. 4.7. The change in momentum then gives the impulse according to Eq. 4.18.

❸ **EXECUTE PLAN** From Eq. 3.15 I see that it takes an object $t = \sqrt{2h/g} = \sqrt{2(1.0\text{ m})/(9.8\text{ m/s}^2)} = 0.45$ s to fall from a height of 1.0 m. Choosing the positive x axis pointing upward and substituting this result into Eq. 3.10, I obtain for the x component of the velocity of the ball just before it hits the floor:

$$v_{x,\text{f}} = 0 + (-9.8\text{ m/s}^2)(0.45\text{ s})^2 = -2.0\text{ m/s}.$$

Now that I know the ball's velocity just before it hits the ground I can obtain the x component of the momentum of the ball just before it hits the ground by multiplying the ball's velocity by its inertia: $p_{x,\text{i}} = (0.20\text{ kg})(-2.0\text{ m/s}) = -0.40\text{ kg}\cdot\text{m/s}$. (I use the subscript i to indicate that this is the initial momentum of the ball before the collision with the floor.) If the ball rebounds with the same speed, then the x component of the momentum after the collision with the floor has the same magnitude but opposite sign: $p_{x,\text{f}} = +0.40\text{ kg}\cdot\text{m/s}$. The change in the ball's momentum is thus

$$\begin{aligned}\Delta p_x &= p_{x,\text{f}} - p_{x,\text{i}} = +0.40\text{ kg}\cdot\text{m/s} - (-0.40\text{ kg}\cdot\text{m/s}) \\ &= +0.80\text{ kg}\cdot\text{m/s}.\end{aligned}$$

The interaction with the ground changes the momentum of the ball, making it rebound. The ball does not constitute an isolated system, and the change in its momentum is due to an impulse delivered by Earth to the ball. To determine the impulse, I substitute the change in momentum of the ball into Eq. 4.18:

$$\vec{J} = \Delta\vec{p} = \Delta p_x\hat{\imath} = (+0.80\text{ kg}\cdot\text{m/s})\,\hat{\imath}.$$ ✔

❹ **EVALUATE RESULT** The x component of the velocity of the ball just before it hits the floor is negative because the ball moves downward, in the negative x direction. After the collision it moves in the opposite direction, and consequently the x components of the changes in velocity and momentum are both positive, as I calculated. Because the x component of the change in the ball's momentum is positive, the impulse is directed upward (in the positive x direction). This makes sense because this impulse changes the direction of travel of the ball from downward to upward.

Because the momentum law takes on the simplest form for an isolated system, we shall usually try to choose an isolated system (unlike the above example). In such a case, we can write out the $\Delta\vec{p}$ term in Eq. 4.17:

$$\vec{p}_f = \vec{p}_i \quad \text{(isolated system)}. \tag{4.19}$$

In this form, the momentum law states that the initial and final values of the momentum of an isolated system are equal. For two colliding (or otherwise interacting) objects moving along the x axis, Eqs. 4.17 and 4.19 become

$$\Delta p_{1x} + \Delta p_{2x} = 0 \quad \text{(isolated system)} \tag{4.20}$$

and

$$p_{1x,f} + p_{2x,f} = p_{1x,i} + p_{2x,i} \quad \text{(isolated system)}. \tag{4.21}$$

Equations 4.20 and 4.21 are equivalent: Eq. 4.20 states that any change in momentum in one object is made up by a change of the same magnitude but in the opposite direction in the other object; Eq. 4.21 states that the sum of the momenta is the same before and after the collision. Which of these two equations you use in solving a problem is a matter of convenience.

4.14 (*a*) Are the changes in velocity in Figure 4.23*a* equal in magnitude? Why or why not? (*b*) Determine the velocity changes of the carts and verify that $m_1/m_2 = -\Delta v_{2x}/\Delta v_{1x}$ in Figure 4.23*a*. (*c*) Determine the initial and final momenta of the two carts. (*d*) What is the momentum of the system before the collision? (*e*) After the collision? (*f*) Are the momentum changes equal in magnitude and opposite in direction? Why or why not?

To convince you of the fundamental importance of the conservation of momentum, let me show you something else. Suppose we let carts 1 and 2 of Figure 4.23 collide with the same initial velocities as before, but this time we put something sticky between them, so that they remain locked together once they hit each other, as in **Figure 4.24**. The velocities of the two carts *after* the collision must now be identical.

4.15 (*a*) Do the two carts in Figure 4.24 still constitute an isolated system like the carts in Figure 4.23? (*b*) What does your answer to part *a* imply about the momentum of the system comprising the two carts after the collision?

As you saw in the preceding checkpoint, the system of the two carts is isolated and so the momentum of the system should not change. **Figure 4.25*a*** on page 140 shows graphically that the final velocity of the combined carts is somewhere between the values of the two initial velocities. Before the collision, cart 1 is at rest and the x component of the velocity of cart 2 is $+0.34$ m/s; after the collision, both carts move at $v_{x,f} = +0.085$ m/s. The x component of the momentum of the system before the collision is thus $p_{1x,i} + p_{2x,i} = 0 + (0.12\ \text{kg})(0.34\ \text{m/s}) = 0.041\ \text{kg}\cdot\text{m/s}$. After the collision, the x component of the momentum of the system is $p_{1x,f} + p_{2x,f} = (0.36\ \text{kg})(0.085\ \text{m/s}) + (0.12\ \text{kg})(0.085\ \text{m/s}) = 0.031\ \text{kg}\cdot\text{m/s} + 0.010\ \text{kg}\cdot\text{m/s} = 0.041\ \text{kg}\cdot\text{m/s}$. The data indeed confirm that the momentum of the system does not change.

Note that the before-collision parts of Figures 4.23*a* and 4.25*a* are identical. The after-collision parts of these two figures are quite different from each other, however. Yet even with these two entirely different outcomes—the carts in Figure 4.23 flying off in opposite directions after colliding and those in Figure 4.25 sticking together—the momentum of the system does not change in both cases:

$$\vec{p}_i = \vec{p}_f \quad \text{(isolated system)}. \tag{4.22}$$

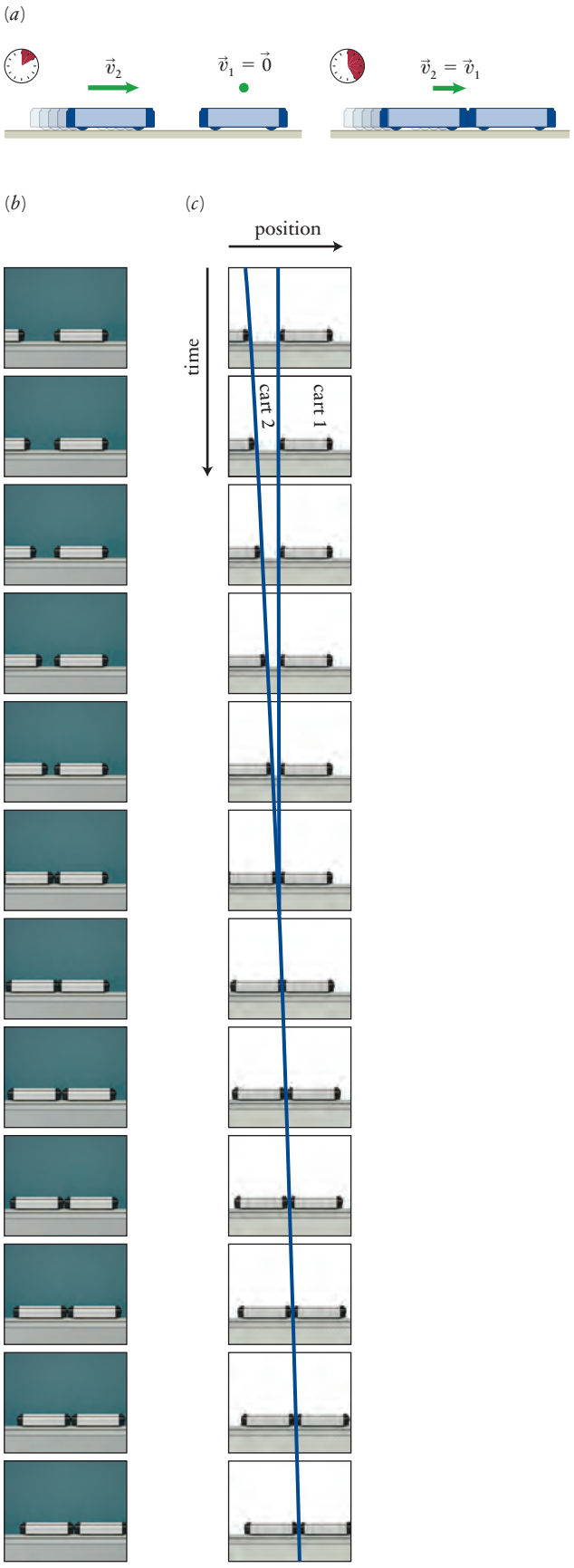
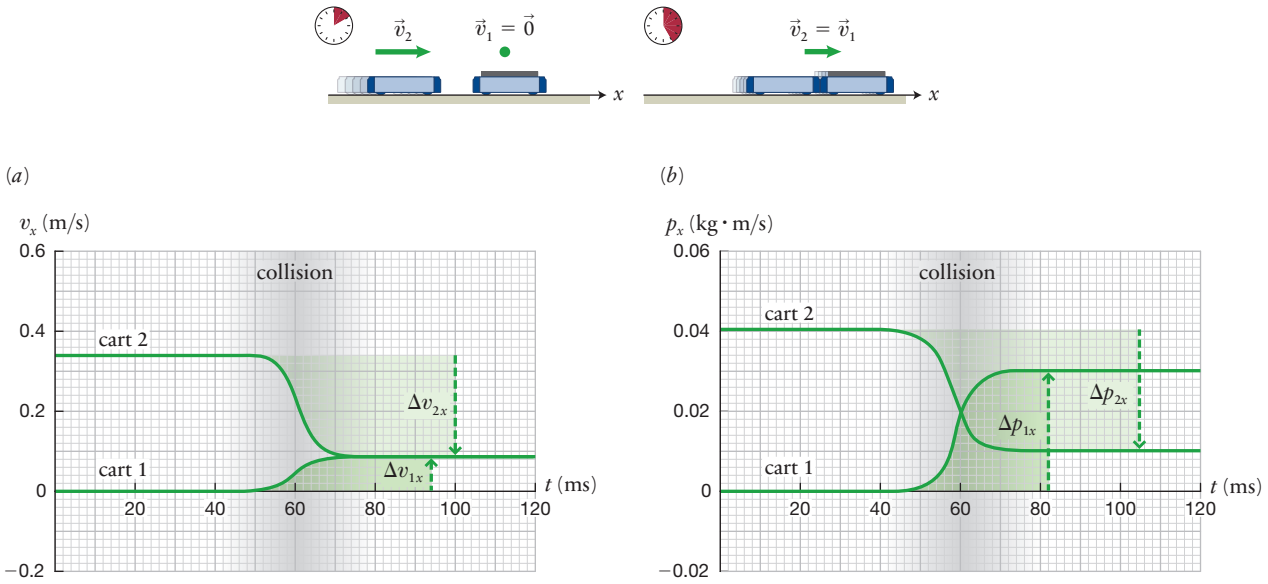


Figure 4.24 (*a*) Two identical carts collide and stick together on a low-friction track. (*b*) High-speed video sequence of such a collision. The length of track visible in each frame is 0.40 m. (*c*) Curves indicating the positions of the rear of cart 1 and the front of cart 2 superimposed on the sequence.

Figure 4.25 (*a*) Velocity-versus-time graph and (*b*) momentum-versus-time graph for the carts of Figure 4.24. Notice that the sum of the momenta is again the same before and after the collision. The inertia of cart 1 is 0.36 kg, and that of cart 2 is 0.12 kg.



The momentum law is not confined to two objects. Suppose we put *three* objects in motion on a low-friction track such that they all end up colliding with one another. Let 1 first collide with 2 and then 2 with 3 (you can pick whatever sequence you like). In the first collision, $\vec{p}_1$ and $\vec{p}_2$ change but their sum remains unchanged. In addition, the sum of all three momenta remains unchanged because nothing has happened to $\vec{p}_3$ yet. When 2 and 3 collide, $\vec{p}_2$ and $\vec{p}_3$ change but the sum $\vec{p}_2 + \vec{p}_3$ does not change. So again, the momentum of the system comprising all three remains constant. In other words, throughout the two collisions, the momentum of the system comprising all three objects, $\vec{p}$, remains unchanged:

$$\vec{p}_{\mathrm{i}} = \vec{p}_{1,\mathrm{i}} + \vec{p}_{2,\mathrm{i}} + \vec{p}_{3,\mathrm{i}} = \vec{p}_{1,\mathrm{f}} + \vec{p}_{2,\mathrm{f}} + \vec{p}_{3,\mathrm{f}} = \vec{p}_{\mathrm{f}} \quad \text{(isolated system)}. \qquad (4.23)$$

This argument can be extended to any number of objects and any number of collisions or interactions.

That we can apply conservation of momentum to predict the final momentum of carts colliding on a low-friction track may seem like nothing special because you are indeed unlikely to encounter this situation outside the physics classroom. Conservation of momentum has much broader applications, however, because it governs *everything that happens in the universe*. Conservation of momentum applies to atoms and elementary particles at the subatomic scale, to stars and galaxies at the cosmic scale, and to everything in between. Next time you fly in an airplane consider this: you are moving forward because of conservation of momentum. The aircraft engines "throw air backward" so the airplane and the passengers move forward. Conservation of momentum is used to solve many science and engineering problems: from calculating the forces of impact during vehicle collisions, to the trajectories of satellites, to the load-carrying capabilities of artificial limbs.

QUANTITATIVE TOOLS

4.16 A railroad car moves at speed v_{i} on a track toward three other railroad cars at rest on the track. All the cars are identical, and each stationary car is some distance from the others. The moving car hits the first stationary car and locks onto it. Then the two hit the second stationary car, lock onto that one, and so on. When all four cars have locked onto one another, what is the speed v_{f} of the four-car train?

Chapter Summary

Inertia (Sections 4.1–4.3, 4.5)

Concepts *Friction* is the resistance to motion that one surface encounters when moving over another surface. In the absence of friction, objects moving along a horizontal track keep moving without slowing down.

Inertia is a measure of an object's tendency to resist a change in its velocity. Inertia is determined entirely by the type of material of which the object is made and by the amount of that material contained in the object. Inertia is related to *mass*, and for this reason we use the symbol m to represent it. The SI unit of inertia is the **kilogram** (kg).

Quantitative Tools If an object of unknown inertia m_u collides with an inertial standard of inertia m_s, the ratio of the inertias is related to the changes in the velocities by

$$\frac{m_u}{m_s} \equiv -\frac{\Delta v_{sx}}{\Delta v_{ux}}. \tag{4.1}$$

Systems and momentum (Sections 4.4, 4.6–4.8)

Concepts A **system** is any object or group of objects that can be separated, in our minds, from the surrounding environment. The **environment** is everything that is not part of the system. You can choose the system however you want, but once you decide to include a certain object in the system, that object must remain a part of the system throughout your analysis.

A system for which there are no external interactions is called an **isolated system.**

An **extensive quantity** is one whose value is proportional to the size or "extent" of the system. An **intensive quantity** is one that does not depend on the extent of the system.

A **system diagram** shows the initial and final conditions of a system.

Quantitative Tools The **momentum** $\vec{p}$ of an object is the product of its inertia and velocity:

$$\vec{p} \equiv m\vec{v}. \tag{4.6}$$

The momentum of a system of objects is the sum of the momenta of its constituents:

$$\vec{p} \equiv \vec{p}_1 + \vec{p}_2 + \cdots. \tag{4.11, 4.23}$$

Conservation of momentum (Sections 4.4, 4.8)

Concepts Any extensive quantity that cannot be created or destroyed is said to be **conserved,** and the amount of any *conserved* quantity in an isolated system is *constant*. Momentum is a conserved quantity, and therefore the momentum of an isolated system is constant. The momentum can be transferred from one object to another in the system, but the momentum of the system cannot change.

Quantitative Tools The momentum of an isolated system is constant:

$$\Delta\vec{p} = \vec{0}. \tag{4.17}$$

Another way to say this is that for an isolated system, the initial momentum is equal to the final momentum:

$$\vec{p}_i = \vec{p}_f. \tag{4.22}$$

The **impulse** $\vec{J}$ delivered to a system is equal to the change in momentum of the system:

$$\vec{J} = \Delta\vec{p}. \tag{4.18}$$

For an isolated system, $\vec{J} = \vec{0}$.

Questions and Problems

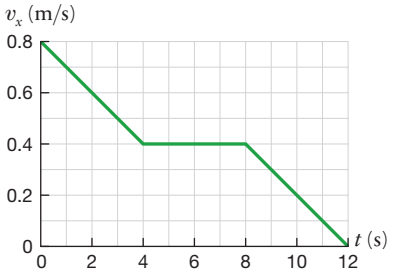
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

4.1 Friction

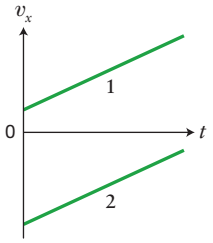
1. Two identical boxes A and B slide over the same rough floor. Box B stops over a distance twice as long as that needed for box A. Assuming a constant frictional force F acting in both cases, compare the initial velocities of the two boxes. •
2. Figure P4.2 shows the velocity of a block of wood as a function of time. The block is sliding over a horizontal surface. Describe the physical processes that led to this graph. •

Figure P4.2



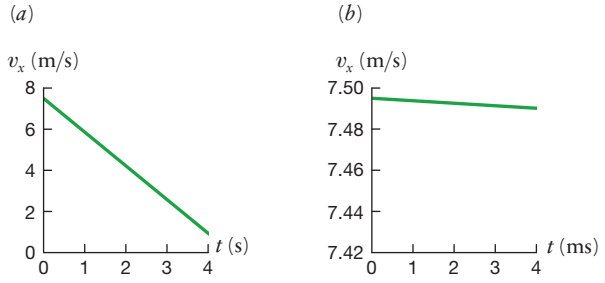
3. The velocity-versus-time graph in Figure P4.3 shows the motion of two different objects sliding across a horizontal surface. Could the change in the x component of velocity with time be attributed to friction in each case? ••

Figure P4.3



4. Consider the two velocity-versus-time graphs shown in Figure P4.4. Are the motions represented by these curves best described as similar or as different? Is the effect of friction on the motion plausibly more pronounced in one case? ••

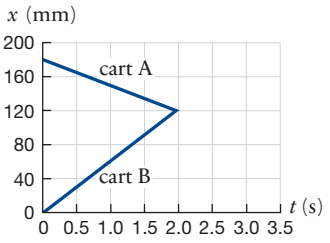
Figure P4.4



4.2 Inertia

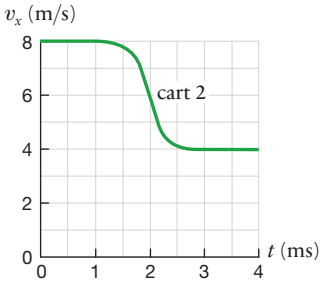
5. Two objects collide on a low-friction track. Object 1 experiences a magnitude of change in velocity $|\Delta\vec{v}_1| = 3$ m/s, while object 2 has $|\Delta\vec{v}_2| = 1$ m/s. How do the inertias of these two objects compare? •
6. In a collision experiment, the ratio of the velocity change between two carts of equal inertia is found to equal 1. What happens to this ratio if the experiment is repeated in the following conditions? (*a*) The inertia of each cart is doubled, and the same initial velocities are used. (*b*) The original value of inertia is used, and the initial velocities of the carts are doubled. •
7. Two carts of equal inertia are moving in opposite directions toward each other. Figure P4.7 represents the positions of the carts until they collide. Sketch the positions of the carts after the collision. •

Figure P4.7



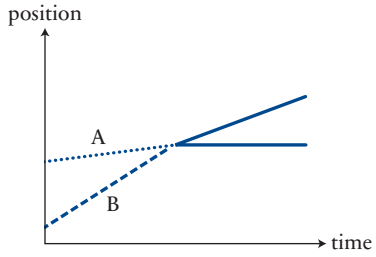
8. Cart 1 initially at rest is struck by cart 2, which has twice the inertia of cart 1. Figure P4.8 shows the velocity of cart 2 as a function of time. Complete the graph by adding the velocity of cart 1 during the same time. ••

Figure P4.8



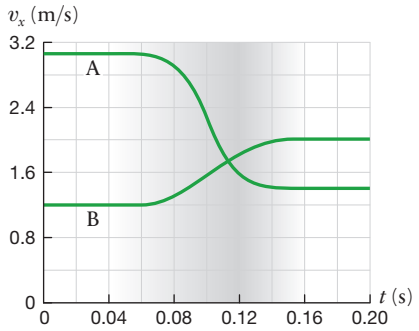
9. Figure P4.9 is the position-versus-time graph for a collision, along a straight line, between two identical amusement-park bumper cars A and B. The inertias of the passengers are different. (*a*) Which post-collision solid line is a continuation of the dotted line for car A? Which is a continuation of the dashed line for car B? (*b*) Which car contains the passenger having the greater inertia? ••

Figure P4.9



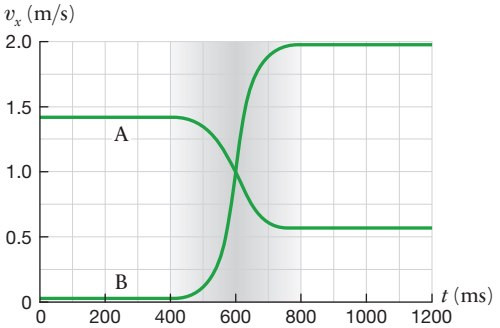
10. Figure P4.10 shows the $v_x(t)$ curves for two carts, A and B, that collide on a low-friction track. What is the ratio of their inertias? ●●

Figure P4.10



11. Figure P4.11 is a velocity-versus-time graph for two objects, A and B, before and after they collide. Object B is initially at rest. What factor relates the inertias of objects A and B? ●●

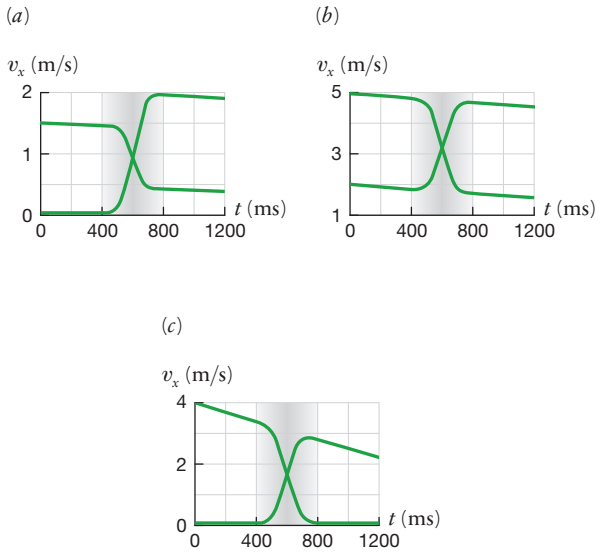
Figure P4.11



4.3 What determines inertia?

12. There are two 3-L bottles, one filled with water and the other filled with milk. Which one will have higher inertia? ●
13. A goldsmith used a 500-g bar of gold and melted it and molded into a small statue of a Roman god. Comment on the change in the inertia of the gold before and after being molded. ●
14. Five different objects are formed from a constant volume V of different materials, and the objects are placed on various surfaces. Rank the following objects in order of increasing inertia: a cubic lead block on ice, a cubic plastic block on ice, a sphere of lead on ice, a pyramid of plastic on concrete, and a pyramid of lead on concrete. ●
15. During a flood-relief operation, two boxes having same inertia, one filled with polythene tent material and the other one filled with rice packets, are dropped from a helicopter. Which one will reach the ground fast? ●●
16. The engine of a moving spaceship suddenly stopped working. What will happen to the motion of the space ship? ●●
17. The velocity-versus-time graphs in Figure P4.17 all depict the motion of carts that have the same size and shape. The different graphs show motion on different tracks: a smooth icy track, a dusty unpolished track, and a rough damaged track. In one case metal carts are used, in one case plastic cars are used, and in one case one cart of each material is used. The carts may initially be moving in opposite directions, in the same direction, or one may be at rest. For each graph, label the track type, cart type, and initial motion. If you cannot determine a property, state why. ●●

Figure P4.17

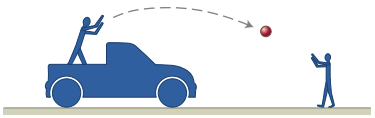


4.4 Systems

18. The boiling point of water on top of a mountain reduced to 80 °C compared to 100 °C at the sea level. Is the boiling point of water intensive or extensive? ●
19. You are riding a bus and thinking about the number of passengers on board. (*a*) Is the number of passengers an extensive or intensive quantity? (*b*) Draw a system diagram to help account for the number of passengers. Where is the system boundary? (*c*) Does the number of passengers remain constant? (*d*) Does the number of passengers remain constant as the bus follows its route while its doors stay closed? ●
20. A quart is a unit of volume. (*a*) If you have 1 quart of water in one container and 1 quart of water in another container and you pour all the water from both into a larger container, what volume of water do you end up with? (*b*) Suppose you have a 1-quart container filled with marbles and a large pail containing 1 quart of water. When you dump the marbles into the pail, is the final volume the same as in part *a*? Why or why not? (*c*) Compare the inertias in each case before and after mixing. ●●

21. Figure P4.21 shows a person on a truck throwing a ball to a friend on the ground. In how many ways can you divide these things into a system and an environment? ●●

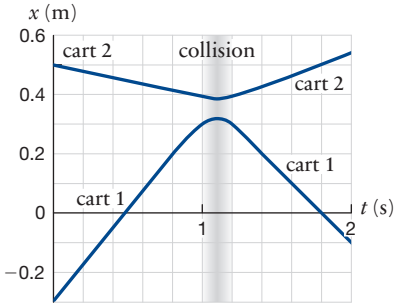
Figure P4.21



4.5 Inertial standard

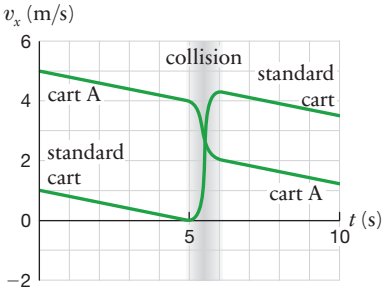
22. If two carts collide on a low-friction track, it is possible for both carts to slow down. Yet Eq. 4.1 $\left(\frac{m_u}{m_s} = -\frac{\Delta v_{sx}}{\Delta v_{ux}}\right)$ seems to say that, because inertia is always positive, if one velocity increases, the other must decrease. Explain the apparent discrepancy. ●
23. Cart A, of inertia 1.0 kg, is initially at rest on a low-friction track; cart B, of unknown inertia, has an initial velocity of +3.0 m/s in your coordinate system. After the two carts collide, the final velocities are $\vec{v}_A = +2.0$ m/s and $\vec{v}_B = -3.0$ m/s. What is the inertia of cart B? ●
24. A 1.0-kg cart is initially at rest on a low-friction track. It is hit by a 3-kg cart moving at 0.2 m/s to the left. After the collision, the 3-kg cart continues to move to the left at 0.1 m/s. What is the velocity of the 1-kg cart? ●●
25. A 1-kg standard cart collides with a 5.0-kg cart initially at rest on a low-friction track. After the collision, the standard cart is at rest and the 5.0-kg cart has a velocity of 0.20 m/s to the left. What was the initial velocity of the standard cart? ●●
26. Figure P4.26 is the position-versus-time graph for a collision between two carts on a low-friction track. Cart 1 has an inertia of 1.0 kg; cart 2 has an inertia of 4.0 kg. (*a*) What are the initial and final velocities of each cart? (*b*) What is the change in the velocity of each cart? (*c*) Do the values you calculated in part *b* satisfy Eq. 4.1? (*d*) Draw a velocity-versus-time graph for this collision. (*e*) Does cart 1 have a nonzero acceleration? If so, when, and what is the sign of the acceleration? (*f*) Does cart 2 have a nonzero acceleration? If so, when, and what is the sign of the acceleration? ●●

Figure P4.26



27. The bumper boats at your local theme park each have an inertia of 90 kg. In boat 1 are a man of unknown inertia, a 45-kg woman, and a 3.0-kg dog. In boat 2 are an 80-kg father, a 50-kg mother, and a 30-kg son. Boat 1 collides at 1.5 m/s with boat 2, which is initially at rest. Two seconds after the collision, boat 2 has moved 2.3 m and boat 1 has moved 0.26 m in the opposite direction. (*a*) Taking the initial direction of motion of boat 1 as positive, what is the velocity of each boat after the collision? (*b*) What is the change in velocity of each boat? (*c*) What is the man's inertia? (*d*) If the collision took 0.50 s, what was each boat's average acceleration during the collision? ●●●
28. A 1-kg standard cart collides with a cart A of unknown inertia. Both carts appear to be rolling with significant wheel friction because their velocities change with time as shown in Figure P4.28. (*a*) What are the carts' velocities at $t = 0$, $t = 5.0$ s, $t = 6.0$ s, and $t = 10$ s? (*b*) When not colliding, are the carts speeding up or slowing down? (*c*) Is the acceleration of each cart the same before and after the collision? (*d*) What is the inertia of cart A? ●●●

Figure P4.28



4.6 Momentum

29. At what speed should a 0.02-kg bullet be shot to have the same momentum as a 0.45-kg soccer ball kicked at 80 km/hr by a professional player? ●
30. The student next to you says, "Momentum is inertia times velocity. That means momentum is proportional to inertia, and so things with more inertia have more momentum." What would you say to the student to clarify the matter? ●
31. A 2000-kg car is moving at the same speed as a 6000-kg bus, as they go on adjacent lanes along the same level road. Both drivers apply their brakes, stopping over the same distance to avoid a red light. Which vehicle will have a bigger change in its momentum? ●
32. A bicyclist coasting along a road at speed v collides with a grasshopper flying in the opposite direction at the same speed. The grasshopper sticks to the bicyclist's helmet. Make a sketch showing (*a*) the two velocity vectors and the two momentum vectors before the collision, (*b*) the velocity and momentum vectors after the collision, and (*c*) the velocity change and the momentum change for the bicyclist and for the grasshopper. ●●
33. Is it possible for the momentum of a system consisting of two carts on a low-friction track to be zero even if both carts are moving? ●●

34. A 1.0-kg standard cart A collides with a 0.10-kg cart B. The x component of the velocity of cart A is +0.60 m/s before the collision and +0.50 m/s after the collision. Cart B is initially traveling toward cart A at 0.40 m/s, and after the collision the x component of its velocity is +0.60 m/s. What are (*a*) the x component of the change in the momentum of A, (*b*) the x component of the change in the momentum of B, and (*c*) the sum of these two x components of the changes in momentum? ••

35. You're rolling solid rubber balls on the kitchen floor. Ball 1 has a density of 1.00×10^3 kg/m^3 and a radius of 25.0 mm. Ball 2 has an unknown density and a radius of 40.0 mm and is initially at rest. You roll ball 1 at an initial speed of 3.00 m/s, and the two balls collide head-on. Ball 1 reverses direction and comes back to you at 2.00 m/s, and after the collision, the speed of ball 2 is 1.00 m/s. With the positive x axis in the direction of ball 1's initial motion, what are (*a*) the magnitudes of the initial and final momenta of ball 1, (*b*) the initial and final momenta of ball 2, and (*c*) the density of ball 2? ••

36. A 1-kg box is accidentally dropped from a third-floor window 7.5 m above the ground. What is the change in its momentum as it strikes the muddy ground, coming to a complete stop? ••

37. From what height would a car have to fall in order for the magnitude of its momentum to equal the magnitude of its momentum when it is moving on a highway at 30 m/s? ••

38. (*a*) Write an expression relating the average acceleration, Δp, and Δt for an object of constant inertia m. (*b*) Given your result in part *a*, what can you say about the magnitude of the acceleration of an object coming to rest after being dropped on a soft bed versus a hard floor? (*c*) Discuss the rationale for airbags in cars. •••

39. You drop a 0.15-kg ball to the floor from a height of 2.0 m, and it bounces to a height of 1.6 m. What is the magnitude of the change in its momentum as a result of the bounce? •••

4.7 Isolated systems

40. When a rifle is fired, it *recoils*, kicking backward into the shooter's shoulder. Why? •

41. A rifle fires a 15-g bullet at 500 m/s and recoils with a speed of 2.5 m/s. What is the inertia of the rifle? •

42. Robert Brown first discovered the random motion of pollen immersed in water through a microscope. When seen with naked eyes, everything seemed to be still, but, upon viewing through a microscope, the random motion of pollen grains was detected. Can you calculate the net momentum of the pollen grains? ••

43. For our solar system, the planets with their moons revolve around the sun with different orbital periods. Even the sun is not static, and it revolves around a spot in the Milky Way galaxy in approximately 250 million years. In that scenario, do you consider our solar system to be an isolated system or not? Justify your answer. ••

44. A car moves with constant velocity on a regular road. Can you say the car is an example of an isolated system? ••

45. In a pairs skating competition, a 75-kg male skater moving at 4.0 m/s collides (gently) with his stationary, 50-kg female partner and raises her in a lift. Neither of them makes a horizontal push at the instant of the pickup. What is the speed of the pair after the collision? (Hint: The vertical motion of the female has no bearing on the question.) ••

46. A girl wearing ice skates stands in a skating rink and throws her backpack to a bench just off the ice. (*a*) If her skates are aligned parallel to the direction of the throw, what happens to her as a result? (*b*) When the backpack lands on the bench, it stops. Does the skater stop at the same instant? If not, where does the momentum of the skater-backpack system go? ••

47. A cart moving on a low-friction track has a momentum of 6 kg · m/s to the right. At the end of the track is a wall. (*a*) What is the momentum of the system that consists of the cart and the wall? (*b*) After the cart collides with the wall, the cart's momentum is 6 kg · m/s to the left. What is the wall's momentum after impact, from the track's perspective? (*c*) Is the system defined in part *a* an isolated system? Why doesn't the wall move? ••

4.8 Conservation of momentum

48. What does conservation of momentum say about the motion of a single cart that doesn't collide with anything as it moves on a low-friction track? •

49. The three carts described in the accompanying table collide simultaneously. What is the momentum of the system of three carts (*a*) before the collision and (*b*) after the collision? (*c*) Is the system of three carts isolated during the collision, according to these data? •

	Momentum p_x (kg · m/s)	
Cart	Before collision	After collision
1	+6.0	−4.0
2	−2.0	+2.0
3	−3.0	+3.0

50. Two carts collide on a track. The initial momentum of cart 1 is +5 kg · m/s $\hat{\imath}$ and its final momentum is −3 kg · m/s $\hat{\imath}$. The initial momentum of cart 2 is −4 kg · m/s $\hat{\imath}$ and its final momentum is +4 kg · m/s $\hat{\imath}$. Would you say that the track is a low-friction track? Explain. •

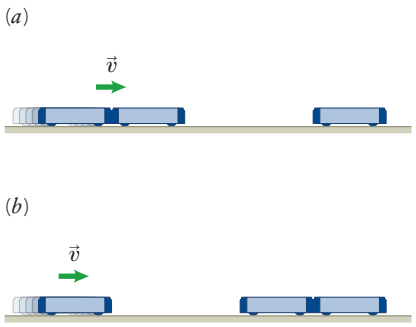
51. Two carts A and B collide on a low-friction track. Measurements show that their initial and final momenta are $\vec{p}_{A,i} = (+10 \text{ kg} \cdot \text{m/s})\,\hat{\imath}$, $\vec{p}_{A,f} = (+2.0 \text{ kg} \cdot \text{m/s})\,\hat{\imath}$, $\vec{p}_{B,i} = (-4.0 \text{ kg} \cdot \text{m/s})\,\hat{\imath}$, and $\vec{p}_{B,f} = (+4.0 \text{ kg} \cdot \text{m/s})\,\hat{\imath}$. (*a*) What is the change in the momentum of each cart during the collision? (*b*) What is the sum of these changes in momenta? (*c*) Is this collision consistent with conservation of momentum? Why or why not? ••

52. Two carts A and B collide on a low-friction track. Measurements show that their initial and final momenta are $\vec{p}_{A,i} = (+3.0 \text{ kg} \cdot \text{m/s})\,\hat{\imath}$, $\vec{p}_{A,f} = (+1.0 \text{ kg} \cdot \text{m/s})\,\hat{\imath}$, $\vec{p}_{B,i} = (+2.0 \text{ kg} \cdot \text{m/s})\,\hat{\imath}$, and $\vec{p}_{B,f} = (-6.0 \text{ kg} \cdot \text{m/s})\,\hat{\imath}$. (*a*) What is the change in the momentum of each cart during the collision? (*b*) What is the sum of these changes in momenta? (*c*) Is this collision consistent with the physics you know? Why or why not? ••

53. A moving object collides with an object at rest. (*a*) Is it possible for both objects to be at rest after the collision? (*b*) Is it possible for just one object to be at rest after the collision? If so, which one? Ignore any effects due to friction. ••

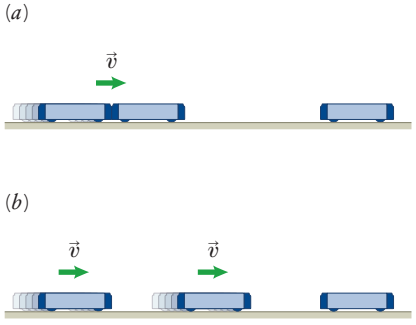
54. Estimate the magnitude of impulse that you impart to a nail (and to whatever the nail is buried in) when you hit it with a hammer. ●●
55. A 1200-kg automobile traveling at 15 m/s collides head-on with a 1600-kg automobile traveling at 10 m/s in the opposite direction. (*a*) Is it possible to predict the velocities of the cars after the collision? (*b*) Is it possible to predict the value that any pertinent physical quantity has immediately after the collision? ●●
56. A 2.0-kg cart and a 3.0-kg cart collide on a low-friction track. The 3.0-kg cart is initially moving at 1.0 m/s to the right, but after the collision it is moving at 5.0 m/s to the right. After the collision, the 2.0-kg cart is moving to the right at 3.0 m/s. (*a*) What was the 2.0-kg cart's initial velocity? (*b*) What would your answer be if, other quantities being unchanged, the 3.0-kg cart were initially moving to the left at 1.0 m/s? ●●
57. Warmly dressed in several layers of clothing, you are standing at the center of a frozen pond. There is not enough friction to permit you to walk to the edge of the pond. How can you save yourself? (Never mind how you got to the center of the pond.) ●●
58. On an icy day, a driver swerves into the opposite lane and gets into a head-on collision with another car whose inertia is twice that of his. Both cars have the same speed when they hit and lock together. What fraction of their original speed do they have when moving together? ●
59. An 80-kg physicist and a friend are ice-skating. The physicist, distracted, collides from behind at 7.0 m/s with his friend, who is skating at 5.0 m/s in the same direction. After the collision, the physicist continues in the same direction at 4.0 m/s, but his friend is now moving (still in the same direction) at 8.0 m/s. What is her inertia? ●●
60. A 1500-kg car is being driven with the top open at 30 m/s $\hat{\imath}$ when a 50-g golf ball moving at -10 m/s $\hat{\imath}$ hits the cushioned seat and stops immediately. What is (*a*) the car's momentum and (*b*) the ball's momentum before the collision? (*c*) What is the change in the car's velocity due to the collision? ●●
61. Sledding down a hill, you are traveling at 10 m/s when you reach the bottom. You (inertia 70 kg) then move across horizontal snow toward a 200-kg boulder but jump off the sled (inertia 5.0 kg) the instant before it hits the boulder. The boulder is sitting on *very* slick ice and moves freely when the sled hits it. The sled bounces back, moving at 6.0 m/s. (*a*) At what speed does the boulder move after the sled hits it? (*b*) If you stayed on the sled, what would your momentum be just before you hit the boulder? (*c*) Suppose you and the sled hit the boulder and continue forward after the collision at 2.0 m/s. Would the boulder's after-collision speed be higher or lower than the speed you calculated in part *a*? ●●
62. A fire hose sprays water against a burning building. The stream of water has cross-sectional area A and density ρ and moves with speed v toward the building. Assume the water splatters against the building without reflecting. What magnitude of impulse does the stream of water deliver during the time interval Δt? ●●
63. Playing pool, you send the cue ball head-on (that is, along the line joining the centers of the two balls) into the stationary 8 ball, and the cue ball stops as a result of the collision. (*a*) Describe the momentum of each ball before and after the collision. (*b*) Show that your answer to part *a* is consistent with Eq. 4.18, $\Delta\vec{p} = \vec{J}$. (Note: Assume that pool balls have identical inertias.) ●●
64. Three identical carts on a low-friction track have putty on their ends so that they stick together when they collide. In Figure P4.64*a*, two carts already stuck together and moving with speed v are about to collide with the third cart initially at rest. In Figure P4.64*b*, the single cart moving with the same speed v is about to collide with the two carts already stuck together and initially at rest. In which case is the final velocity of the trio of carts greater? ●●

Figure P4.64



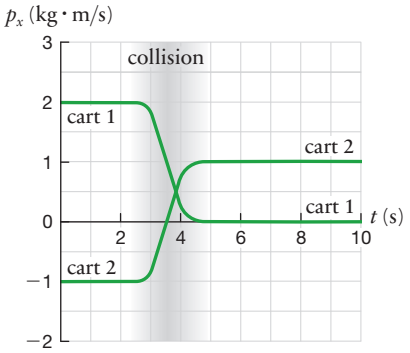
65. A 20-kg mail cart rolls on a smooth floor at 5 m/s when a 2-kg package drops from a shelf directly above and comes to rest in the cart. What is the change in the cart's speed? ●●
66. Three identical carts on a low-friction track have putty on their ends so that they stick together when they collide. In Figure P4.66*a*, two carts already stuck together and moving at speed v collide with the single cart initially at rest. In Figure P4.66*b*, there are two collisions: First the center cart, moving at speed v, collides with the cart on the right (and they stick together). Then the left cart, moving at speed v, collides with those two carts. In which case is the final speed of the three carts greater? ●●

Figure P4.66



67. Figure P4.67 is a momentum-versus-time graph for a collision between two carts. Carts 1 and 2 have inertias of 1.0 kg and 3.0 kg, respectively. What are the x components of the initial and final velocities of (*a*) cart 1 and (*b*) cart 2? (*c*) What are the momentum changes $\Delta\vec{p}_1$ and $\Delta\vec{p}_2$? (*d*) Do the values you calculated in part *c* satisfy Eq. 4.18, $\Delta\vec{p} = \vec{J}$? (*e*) Draw the velocity-versus-time graph for this collision. (Hint: You can do this by making some simple changes to Figure P4.67.) •••

Figure P4.67

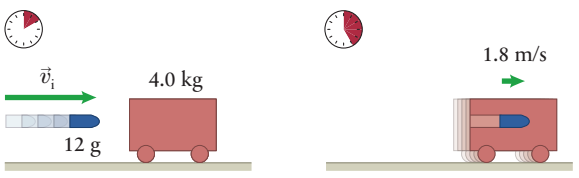


68. Carts A and B collide on a low-friction track. Rank, from largest to smallest, the following four collisions according to the magnitude of the change in the momentum of cart B, which has twice the inertia of cart A. (*a*) A initially moving right at 1.0 m/s, B initially stationary; stick together on impact. (*b*) A initially stationary, B initially moving left at 1.0 m/s; stick together on impact. (*c*) A initially moving right at 1.0 m/s; after impact, A moving left at 0.33 m/s, B moving right at 0.67 m/s. (*d*) A initially moving right at 1.0 m/s, B initially moving left at 1.0 m/s; stick together on impact. •••
69. A red cart and a green cart, both traveling at speed v on a low-friction track, approach each other head-on. The red cart is moving to the right; the green cart, whose inertia is three times that of the red cart, is moving to the left. A black cart that has a spring on its left end and putty on its right end sits on the track between the two moving carts. The moving carts hit the black cart at the same instant. After the collision, the red cart moves left at speed v, and the green and black carts move together at speed $v/5$. What is the inertia of the black cart, expressed as a multiple of the inertia of the red cart? •••
70. A 1.0-kg standard cart collides on a low-friction track with cart A. The standard cart has an initial x component of velocity of +0.40 m/s, and cart A is initially at rest. After the collision the x component of velocity of the standard cart is +0.20 m/s and the x component of velocity of cart A is +0.60 m/s. After the collision, cart A continues to the end of the track and rebounds with its speed unchanged. Before the carts collide again, you drop a lump of putty onto cart A, where it sticks. After the second collision, the x component of velocity of the standard cart is −0.20 m/s and the x component of velocity of cart A is +0.40 m/s. What is the inertia of the putty? •••
71. A 400-kg shipboard cannon fires a 20-kg ball at 60 m/s. The cannon's resulting recoil speed across the deck is regarded as excessive. How much inertia (in the form of sandbags, say) must be added to the cannon if its recoil speed must be reduced to 2.0 m/s? •••
72. You point a tennis-ball serving machine straight up and put a piece of plastic wrap across the top of the barrel, placing a 10-g marble on top of the barrel. When a 58-g tennis ball in the machine is fired, it sends the marble into the air. The marble rises 200 m above the barrel mouth, and the tennis ball rises 54 m. Neglecting the plastic wrap, what is the ball's launch speed? •••

Additional Problems

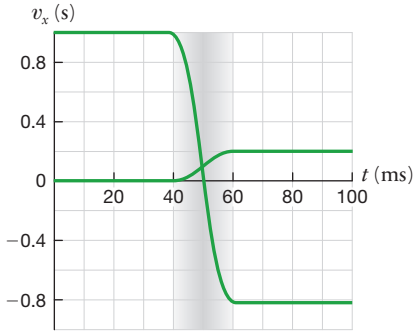
73. The momentum ratio of a cannon grapeshot and rifle bullet is 3:100. If the mass of the grapeshot is 5 kg and the mass of the bullet is 0.05 kg, find the relation between their velocities. •
74. A 30,000-kg trailer truck approaches a one-lane bridge at 2.2 m/s; a 2400-kg minivan approaches the bridge from the other direction at 30 m/s. Which vehicle has a momentum of greater magnitude? •
75. You and a friend are bowling. She rolls her 4.5-kg ball at 10 m/s; you roll your ball at 8.0 m/s. What inertia must your ball have if its momentum is to be the same as hers? •
76. Two identical carts collide in three experiments on a low-friction track. In each case, cart A initially has velocity $\vec{v}$ and cart B is at rest. Sketch the velocity-versus-time graph for each cart (*a*) if the carts stick together after the collision, (*b*) if cart A is stationary after the collision, and (*c*) if cart A has velocity $-\vec{v}/10$ after the collision. ••
77. Draw diagrams that show the initial and final velocity vectors and the initial and final momentum vectors when a rapidly moving golf ball hits (*a*) a golf ball at rest and (*b*) a basketball at rest. In each case, assume that the golf ball moves along the line connecting the centers of the two balls. ••
78. The position of a certain airplane during takeoff is given by $x = \frac{1}{2}bt^2$, where $b = 2.0\ \text{m/s}^2$ and $t = 0$ corresponds to the instant at which the airplane's brakes are released at position $x = 0$. The empty but fueled airplane has an inertia of 35,000 kg, and the 150 people on board have an average inertia of 65 kg. What is the magnitude of the airplane's momentum 15 s after the brakes are released? ••
79. A bullet of inertia m_{bullet} is fired horizontally at a speed $v_{\text{bullet,i}}$ into a stationary wooden block of inertia m_{block} lying on a low-friction surface. The bullet passes through the block and emerges with a speed $v_{\text{bullet,f}}$. Determine the final speed of the block in terms of the given quantities. ••
80. The speed of a bullet can be measured by firing it at a wooden cart initially at rest and measuring the speed of the cart with the bullet embedded in it. Figure P4.80 shows a 12-g bullet fired at a 4.0-kg cart. After the collision, the cart rolls at 1.8 m/s. What is the bullet's speed before it strikes the cart? ••

Figure P4.80



81. The World War II–era rocket launcher called the bazooka was essentially a tube open at both ends. On the basis of momentum considerations, how is the firing of a bazooka different from the firing of a cannon, which is a tube open at one end only? ●●
82. In a football game, a 95-kg player carrying the ball can run the 50-m dash in 5.5 s. Two opponents are available to stop him with a head-on tackle. One opponent has an inertia of 110 kg and a 50-m time of 6.6 s; the other has an inertia of 90 kg and a 50-m time of 5.1 s. Which has the better chance of stopping the 95-kg player, and why? ●●
83. In the days before rocketry, some people argued that rocket engines would not work in space because there is no atmosphere for the exhaust to push against. Even today, some people think that a rocket requires a launch pad to push against in order to lift off. Refute such arguments with an argument based on momentum considerations. ●●
84. In some collisions, the velocity of one participant changes little while that of the other changes a lot, as Figure P4.84 illustrates. (*a*) In which direction (positive or negative) are the objects moving before the collision? (*b*) After the collision? (*c*) What is the ratio of the inertia of the larger object to the inertia of the smaller object? (*d*) Does friction play an important role in this collision? ●●

Figure P4.84



85. When a truck and a car both moving with same velocities collide head on, the car gets the maximum damage. Justify. ●●
86. A single-stage rocket is in deep space coasting at $v_{\text{rocket,i}} = 2.0 \times 10^3$ m/s. It fires its engine, which has an exhaust speed of $v_{\text{exhaust}} = 1.0 \times 10^3$ m/s. What is the rocket's speed $v_{\text{rocket,f}}$ after it has lost one-third of its inertia by exhausting burned fuel? Assume that the fuel is ejected all at once. ●●
87. In the process of moving out of your house, you are dropping stuff out a second-floor window to a friend 4.0 m below. You are about to drop a 6.0-kg stereo speaker when you begin to worry that catching anything that has a momentum greater than 50 kg · m/s might be harmful. ●●● CR
88. You are in mission control in Florida, supervising a probe on its way to Mars. The most recent report indicates that the probe will soon need to fire a quick burst of its booster rocket to increase its speed by 5.2 m/s. You know that the spent fuel is ejected at a speed $v_{\text{fuel}} = 800 \text{ m/s} - v_i$, where v_i is the initial speed of the probe before firing, and that 10% of the probe's inertia is currently fuel. Your main worry is whether enough fuel is left for landing maneuvers. ●●● CR
89. You have just launched your new hot-air balloon and are hovering 30.0 m above the ground. As you are checking equipment, the burner fails. You know that if you do nothing, the balloon will descend with a constant acceleration of 1.50 m/s^2, and that the balloon's basket will be seriously damaged if it hits the ground with a momentum of 2850 kg · m/s or more. The combined inertia of the balloon, basket, and all equipment and fuel is 195 kg, not counting yourself (85.0 kg) and ten 5.00-kg sandbags carried as ballast. You make some quick calculations and take action to save the day. ●●● CR
90. A single gas molecule of inertia m is trapped in a box and travels back and forth with constant speed v between opposite walls A and B a distance ℓ apart. At each collision with a wall, the molecule reverses direction without changing speed. Write algebraic expressions for (*a*) the magnitude of the change in momentum of the molecule as it collides with wall B, (*b*) the amount of time that elapses between collisions with wall B, (*c*) the number of collisions per second the molecule makes with wall B, and (*d*) the change in momentum undergone by wall B, per second, as a result of these collisions. ●●●
91. Consider a two-stage rocket made up of two engine stages, each of inertia m when empty, and a payload of inertia m. Stages 1 and 2 each contain fuel of inertia m, so that the rocket's inertia before any fuel is spent is $5m$. Each stage exhausts fuel at a speed $v_{\text{fuel}} = v_{\text{ex}} - v_{\text{i}}$, where v_{ex} is the speed of the exhaust relative to the rocket and v_{i} is the speed of the rocket relative to the ground at the instant the fuel is spent. The rocket is initially at rest in deep space. Stage 1 fires, ejects its fuel all at once, and then detaches from the remainder of the rocket. The same process is repeated with stage 2. (*a*) What is the final speed of the payload? (Hint: Determine the speed of the rocket after stage 1 fires. Then, because stage 1 is detached from the rocket, *redefine the system* to include only the payload and stage 2, and determine the speed of the rocket after stage 2 fires.) (*b*) Now consider another rocket of inertia $5m$ but with only a single engine stage, of inertia $2m$, carrying fuel of inertia $2m$. What is the final speed of the payload in this case? (*c*) Which design yields a higher payload speed: two stage or single stage? Why? ●●●



5

Energy

5.1 Classification of collisions
5.2 Kinetic energy
5.3 Internal energy
5.4 Closed systems

5.5 Elastic collisions
5.6 Inelastic collisions
5.7 Conservation of energy
5.8 Explosive separations

CONCEPTS

QUANTITATIVE TOOLS

Now that we know about conservation of momentum, can we determine the final velocities of two colliding objects if the only things we know are the initial velocities and the fact that the momentum is conserved? The answer is *no* if our only tool is the momentum law. In the collisions represented in Figures 4.23 and 4.25, for instance, momentum is unchanged in both cases, and yet the two outcomes are definitely not the same. So knowing only that momentum remains constant isn't enough. We need additional information in order to predict future positions and velocities. In the process of looking for this additional information, we shall develop another conservation law—the law of conservation of *energy*.

5.1 Classification of collisions

If you look at the velocity-versus-time graphs in Chapter 4, you will notice that the velocity difference in the two carts, $\vec{v}_2 - \vec{v}_1$, in most cases has the same magnitude before and after the collision. **Figure 5.1** shows two graphs of collisions we considered in Chapter 4 and highlights the velocity difference before and after the collision. This difference is the **relative velocity** of the carts: $\vec{v}_{12} \equiv \vec{v}_2 - \vec{v}_1$ is the velocity of cart 2 relative to cart 1 and $\vec{v}_{21} \equiv \vec{v}_1 - \vec{v}_2$ is the velocity of cart 1 relative to cart 2. The subscript in the relative velocity symbol is always shown with the object we are studying printed last: v_{12} is the velocity *of cart 2* and v_{21} is the velocity *of cart 1*. (We cannot use Δ here because we reserve that symbol to denote the change in a *single* quantity; now we are dealing with the difference between the same quantity for two different objects.)

The magnitude of the relative velocity is called the **relative speed.** Thus $v_{12} \equiv |\vec{v}_2 - \vec{v}_1|$ is the speed of cart 2 (last subscript in v_{12}) relative to cart 1. For motion along the x axis, the relative speed is the absolute value of the difference in the x components of the velocities: $v_{12} \equiv |v_{2x} - v_{1x}|$. Note that the sequence of the subscripts for relative speeds does not matter: $v_{12} \equiv |v_{2x} - v_{1x}| = |v_{1x} - v_{2x}| = v_{21}$ (the speed of cart 2 relative to cart 1 is equal to the speed of cart 1 relative to cart 2).

A collision in which the relative speed before the collision is the same as the relative speed after the collision is called an **elastic collision.** Collisions between hard objects are, for the most part, elastic collisions. For instance, a superball bouncing off a hard floor bounces up with nearly the same speed with which it came down. Thus the relative speed of floor and ball does not change, and the collision is elastic.

5.1 Two cars are moving along a highway with neither one accelerating. Is their relative speed equal to the difference between their speeds? Why or why not?

Any collision for which the relative speed after the collision is lower than that before the collision is called an **inelastic collision.** If you drop a tennis ball to the floor, it rebounds with a lower speed than it had when it hit the floor. The relative speed is reduced, and the collision is inelastic.

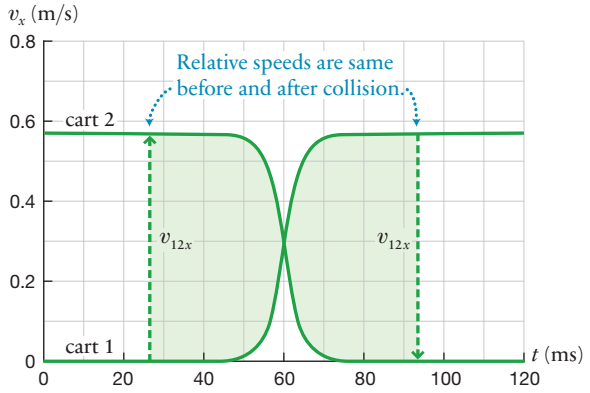
A special type of inelastic collision is one in which the two objects move together after the collision so that their relative speed is reduced to zero. We call this special case a **totally inelastic collision.** Imagine dropping a ball of dough to the floor. Splat! The dough sticks to the floor, and the relative speed of dough and floor is reduced to zero. The collision in Figure 4.25 is another example of an inelastic collision.

As you might imagine, whether a collision is elastic, inelastic, or totally inelastic depends on the properties of the objects involved. However, if we know what happens to the relative speed in a collision, we can use that knowledge together with conservation of momentum to determine the final velocities.

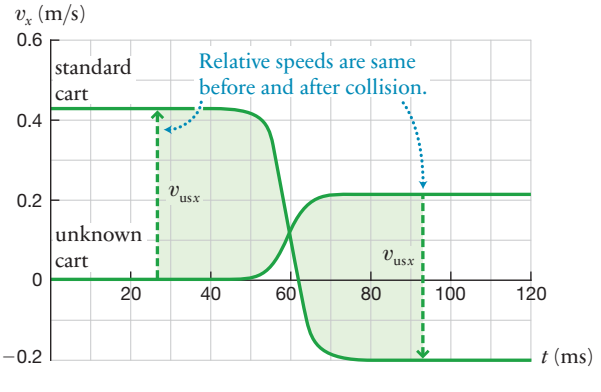
CONCEPTS

Figure 5.1 Velocity-versus-time graphs of (*a*) two identical carts colliding on a low-friction track, with one of the carts initially at rest, and (*b*) a standard cart colliding with a cart of unknown inertia that is initially at rest. Notice that for both collisions the relative speeds of the two carts are the same before and after the collision.

(*a*) Moving cart collides with identical cart at rest



(*b*) Standard cart collides with cart of unknown inertia at rest



EXERCISE 5.1 Classifying collisions

Are the following collisions (for motion along the x axis) elastic, inelastic, or totally inelastic? (*a*) A red billiard ball moving at $v_{rx,i} = +2.2$ m/s hits a white billiard ball initially at rest. After the collision, the red ball is at rest and the white ball moves at $v_{wx,f} = +1.9$ m/s. (*b*) Cart 1 moving along a track at $v_{1x,i} = +1.2$ m/s hits cart 2 initially at rest. After the collision, the two carts move at $v_{1x,f} = +0.4$ m/s and $v_{2x,f} = +1.6$ m/s. (*c*) A piece of putty moving at $v_{px,i} = +22$ m/s hits a wooden block moving at $v_{bx,i} = +1.0$ m/s. After the collision, the two move at $v_{x,f} = +1.7$ m/s.

SOLUTION (*a*) The initial relative speed is $v_{wr,i} = |v_{rx,i} - v_{wx,i}| = |+2.2 \text{ m/s} - 0| = 2.2$ m/s; the final relative speed is $v_{wr,f} = |v_{rx,f} - v_{wx,f}| = |0 - 1.9 \text{ m/s}| = 1.9$ m/s, lower than the initial relative speed, which means the collision is inelastic. ✔

(*b*) $v_{12i} = |v_{2x,i} - v_{1x,i}| = |0 - (+1.2 \text{ m/s})| = 1.2$ m/s; $v_{12f} = |v_{2x,f} - v_{1x,f}| = |+1.6 \text{ m/s} - (+0.4 \text{ m/s})| = 1.2$ m/s. Because the relative speeds are the same, the collision is elastic. ✔

(*c*) After the collision, both the putty and the block travel at the same velocity, making their relative speed zero. The collision is totally inelastic. ✔

5.2 (*a*) An outfielder catches a baseball. Is the collision between ball and glove elastic, inelastic, or totally inelastic? (*b*) When a moving steel ball 1 collides head-on with a steel ball 2 at rest, ball 1 comes to rest and ball 2 moves away at the initial speed of ball 1. Which type of collision is this? (*c*) Is the sum of the momenta of the two colliding objects constant in part *a*? In part *b*?

5.2 Kinetic energy

Relative speed is not an extensive quantity, and so we cannot develop an accounting scheme for it as we did for momentum in Chapter 4. The trick in studying elastic collisions therefore is to obtain a quantity that allows us to express the fact that the relative speed doesn't change in the form: (something of object 1) + (something of object 2) doesn't change. As I shall show you in Section 5.5, for an object of inertia m moving at speed v this something is the quantity $K = \frac{1}{2}mv^2$, called the object's **kinetic energy** (literally "energy* associated with motion"). Unlike relative speed, which refers to two objects, kinetic energy is associated with a single object. Furthermore, kinetic energy is always positive and independent of the direction of motion (which means that it is a scalar). A ball moving at speed v to the left has exactly the same kinetic energy as the same ball moving at speed v to the right (or any other direction in space). Whenever the speed of an object changes, the kinetic energy changes.

5.3 Is kinetic energy an extensive quantity?

To develop some feel for this new quantity and its relationship to relative speed, let us calculate the kinetic energies of the carts before and after the collisions shown in **Figure 5.2**. The two collisions have identical initial conditions, but one is elastic (Figure 5.2*a*) and the other is totally inelastic (Figure 5.2*b*).

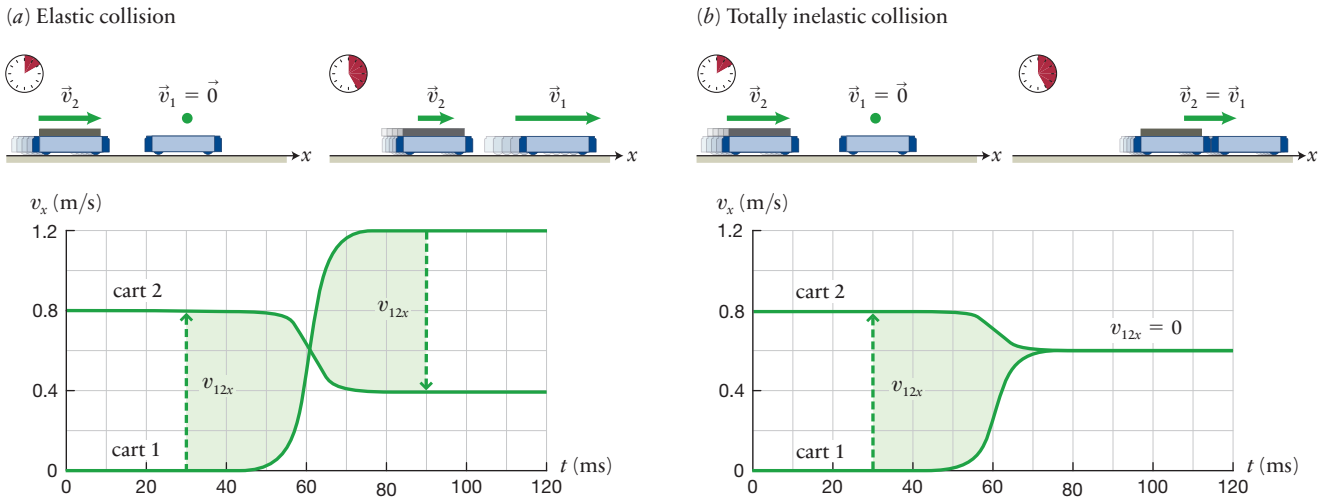
Table 5.1 gives the initial and final kinetic energies of the carts, obtained by reading off the x components of the velocities from the figure. For the elastic collision, where the relative speed doesn't change, the sum of the two kinetic energies before the collision is equal to the sum after the collision: $(0.12 + 0)\ \text{kg}\cdot\text{m}^2/\text{s}^2 = (0.086 + 0.029)\ \text{kg}\cdot\text{m}^2/\text{s}^2$. For the totally inelastic collision, both the relative speed and the sum of the two kinetic energies change.

In general we observe:

In an elastic collision, the sum of the kinetic energies of the objects before the collision is the same as the sum of the kinetic energies after the collision.

*For now think of "energy" as the capacity to do things like accelerating and heating objects. We shall develop a more complete picture of what energy is in later sections and chapters.

Figure 5.2 Velocities for elastic and inelastic collisions between two carts. Cart 1 has inertia 0.12 kg, and cart 2 has inertia 0.36 kg.



CONCEPTS

Table 5.1 Kinetic energy in elastic and totally inelastic collisions

	Inertia m (kg)	ELASTIC Velocity v_x (m/s)		ELASTIC Kinetic energy $\frac{1}{2}mv^2$ (kg·m²/s²)		TOTALLY INELASTIC Velocity v_x (m/s)		TOTALLY INELASTIC Kinetic energy $\frac{1}{2}mv^2$ (kg·m²/s²)	
		before	*after*	*before*	*after*	*before*	*after*	*before*	*after*
Cart 1	0.12	0	+1.2	0	0.086	0	+0.60	0	0.022
Cart 2	0.36	+0.80	+0.40	0.12	0.029	+0.80	+0.60	0.12	0.065
Relative speed		**0.80**	**0.80**			**0.80**	**0**		
Kinetic energy of system				**0.12**	**0.12**			**0.12**	**0.086**

In Section 5.5 we'll see that this statement is equivalent to the requirement that the relative speed remains the same.

For problems involving elastic collisions, we now have two tools: the momentum conservation law, which tells us that the momentum of an isolated system doesn't change, and the fact that the relative speed of the colliding objects doesn't change (or, alternatively, the fact that the kinetic energy of the system doesn't change). Let's now use both tools in an example.

EXAMPLE 5.2 Carts colliding

(*a*) Is the collision in Figure 4.23 elastic, inelastic, or totally inelastic? How can you tell? (*b*) Verify your answer by comparing the initial kinetic energy of the two-cart system with the final kinetic energy.

1 GETTING STARTED In Figure 4.23, it looks as if the initial and final relative speeds are the same, which makes the collision elastic. The problem asks me to confirm this fact by calculating the initial and final kinetic energies of the system.

2 DEVISE PLAN To answer part *a*, I need to determine the initial and final relative speeds of the carts from the velocities, which I get from Figure 4.23*a*: $v_{1x,i} = 0$; $v_{2x,i} = +0.34$ m/s; $v_{1x,f} = +0.17$ m/s; $v_{2x,f} = -0.17$ m/s. To answer part *b*, I use $K = \frac{1}{2}mv^2$. The inertias of the carts are given in the figure caption: $m_1 = 0.36$ kg and $m_2 = 0.12$ kg.

3 EXECUTE PLAN (*a*) $v_{12i} = |v_{2x,i} - v_{1x,i}| = |(+0.34\text{ m/s}) - 0| = 0.34$ m/s; $v_{12f} = |v_{2x,f} - v_{1x,f}| = |(-0.17\text{ m/s}) - (+0.17\text{ m/s})| = 0.34$ m/s. The relative speed is unchanged, and thus the collision is elastic. ✔

(*b*) The initial values are

$$K_{1i} = \tfrac{1}{2}m_1v_{1i}^2 = \tfrac{1}{2}(0.36\text{ kg})(0)^2 = 0$$

$$K_{2i} = \tfrac{1}{2}m_2v_{2i}^2 = \tfrac{1}{2}(0.12\text{ kg})(0.34\text{ m/s})^2 = 0.0069\text{ kg·m}^2/\text{s}^2$$

so

$$K_i = K_{1i} + K_{2i} = 0.0069\text{ kg·m}^2/\text{s}^2.$$

The final values are

$$K_{1f} = \tfrac{1}{2}m_1v_{1f}^2 = \tfrac{1}{2}(0.36\text{ kg})(0.17\text{ m/s})^2 = 0.0052\text{ kg·m}^2/\text{s}^2$$

$$K_{2f} = \tfrac{1}{2}m_2v_{2f}^2 = \tfrac{1}{2}(0.12\text{ kg})(-0.17\text{ m/s})^2 = 0.0017\text{ kg·m}^2/\text{s}^2$$

so $K_f = (0.0052\text{ kg·m}^2/\text{s}^2) + (0.0017\text{ kg·m}^2/\text{s}^2)$
$= 0.0069\text{ kg·m}^2/\text{s}^2$,

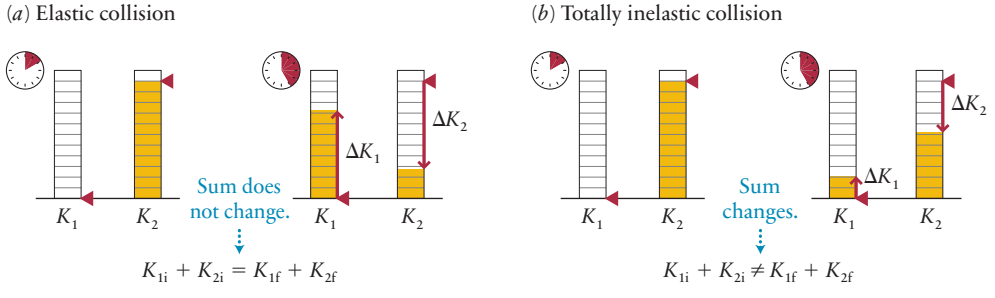
which is the same as before the collision, as it should be for an elastic collision. ✔

4 EVALUATE RESULT Because I've reached the same conclusion—the collision is elastic—using two approaches, I can be pretty confident that my solution is correct.

CONCEPTS

Because kinetic energy is a scalar extensive quantity, bar diagrams are a good way to visually represent changes in this quantity. **Figure 5.3**, for example, shows the initial and final kinetic energies of the carts in the two collisions of Figure 5.2. Before the collision, only cart 2 has kinetic energy; after the collision, both carts have kinetic energy. For the elastic collision (Figure 5.3*a*), the

Figure 5.3 Bar diagrams representing the initial and final kinetic energies of the carts for the collisions shown in Figure 5.2. (*a*) The sum of the kinetic energies does not change for the elastic collision because the changes are equal and opposite. (*b*) The sum of the kinetic energies changes for the totally inelastic collision: The change in K_1 is small but positive; the change in K_2 is greater but negative.



sum of the kinetic energies after the collision is the same as the sum before the collision because the change in the kinetic energy of cart 1 is the negative of that of cart 2. For the totally inelastic collision (Figure 5.3*b*), the changes in the kinetic energies do not cancel, and so the system's kinetic energy after the collision is not equal to the system's kinetic energy before the collision.

5.4 A moving cart collides with an identical cart initially at rest on a low-friction track, and the two lock together. What fraction of the initial kinetic energy of the system remains in this totally inelastic collision?

Figure 5.4 Before or after? (*a*) Inelastic collisions are an irreversible process. There is no question that the picture was taken *after* the accident because the state of the two cars has changed. (*b*) Elastic collisions are a reversible process. Can you tell whether this picture was taken before or after the ball collided with the racquet?

(*a*) (*b*)





5.3 Internal energy

In both inelastic and totally inelastic collisions, the relative speed changes and therefore the kinetic energy of the system changes. For example, the decrease in the kinetic energy of cart 2 in Figure 5.2*b*, represented by the downward pointing arrow labeled ΔK_2 in Figure 5.3*b*, is larger than the increase in the kinetic energy K_1 of cart 1, resulting in a decrease in the kinetic energy of the stuck-together carts. Where does this kinetic energy go? Does it simply vanish, or does it go elsewhere?

We can determine the answer to this question by looking more closely at inelastic collisions between objects and noticing that changes occur in the **state** of one or both objects. What I mean by *state* is the condition of the object as specified by some complete set of physical parameters: shape, temperature, whatever—*every possible physical variable that defines the object*. In inelastic collisions, objects deform (their shape changes) and heat up (their temperature changes): A ball of dough changes shape as it hits the ground, two cars deform as they collide, hands vigorously clapped together heat up and make sound.

The transformation of a system from an initial state to a final state is called a **process.** Processes cause change (see Section 1.4), and so in physics we aim to understand processes. Collisions that change either the motion or the state of objects are an example of a process. Other examples are the melting of an ice cube, the burning of fuel, the flow of a liquid, and an explosive separation. Initially we focus on collisions because they are easily visualized and we can study what happens in collisions using carts on tracks. As we gain experience, we shall replace collisions with other processes.

In inelastic collisions, the state of the objects after the collision is different from the state before the collision. If I were to make a movie of the ball of dough hitting the ground and then play the movie in reverse, you'd have no trouble telling that the movie was being played in the wrong direction. This is because this inelastic collision involves changes that cannot undo themselves—the dough is flat and at rest after the collision and round and moving before (and never the other way around). The same is true for a collision between two cars: The cars are damaged after the collision and not before, and it is not possible to repair them simply by pulling them apart (**Figure 5.4*a***). After the clapping, your hands are warmer than before, and you cannot cool your hands by "unclapping" them. All these inelastic collisions are **irreversible processes,** which means that the changes that occur in the state of the colliding objects cannot spontaneously undo themselves.

In contrast, viewing a movie of a superball bouncing off the floor or two carts colliding on a low-friction track (without sticking together), you'd be hard pressed to tell whether the movie is playing forward or in reverse. This is because elastic collisions are a **reversible process,** which means there are no permanent changes in the state of the colliding objects. The objects look the same before and after the collision (Figure 5.4*b*).

I have summarized these facts in **Table 5.2**. Notice how a change in the relative speed (and therefore a change in the sum of the kinetic energies) goes hand in hand with a change in the state: The sum of the kinetic energies of two colliding objects doesn't change unless their states change. Let us look at this connection with an eye to formulating a new conservation law.

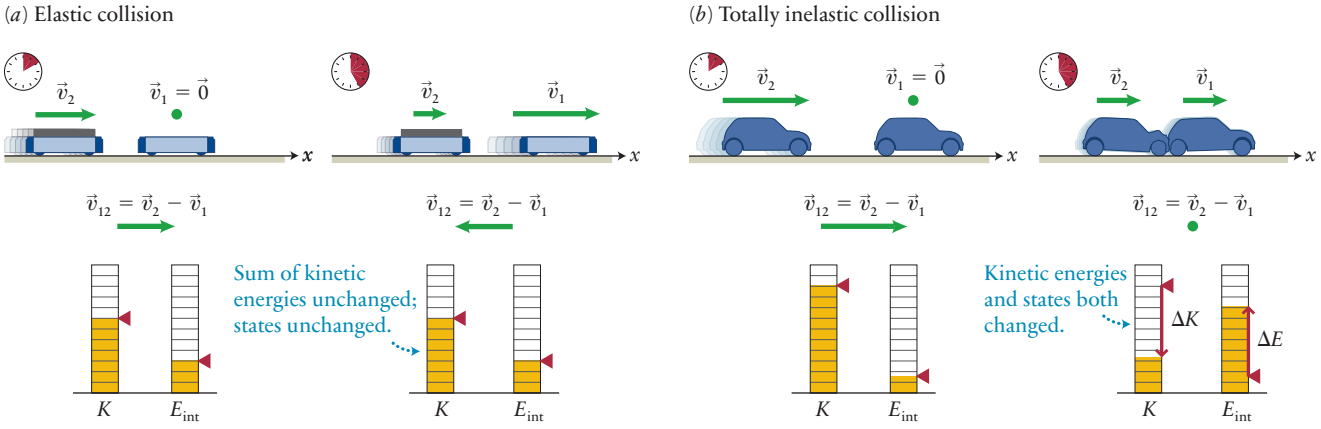
Suppose we could associate some quantity having the same units as kinetic energy ($\mathrm{kg \cdot m^2/s^2}$) with the state of an object—let's call this quantity the object's **internal energy** (denoted by E_{int}). Suppose further that we could arrange things in such a way that in an inelastic collision the increase in the sum of the internal energies of the colliding objects is equal to the decrease in their kinetic energies. This would mean that in an inelastic collision one form of energy is converted to another form (kinetic to internal) but the sum of the kinetic and internal energies—collectively called the **energy** of the system—doesn't change.

Table 5.2 Elastic and inelastic collisions

Collision type	Relative speed	State
elastic	unchanged	unchanged
inelastic	changed	changed
totally inelastic	changed (becomes zero)	changed

CONCEPTS

Figure 5.5 (*a*) Elastic collision. Because the relative speed of the carts does not change, the sum of the kinetic energies is also unchanged. The carts' states do not change, and so their internal energies are also unchanged. (*b*) Inelastic collision. Because the relative speed changes as a result of the collision, the kinetic energy changes. The cars' states also change, and so the internal energy changes too. For both collisions the sum of the kinetic and internal energies is constant.



In Chapter 7 we'll learn how to specify the state of an object and how to calculate the corresponding internal energy. For now, all we need to do is account for the missing kinetic energy in an inelastic collision by saying:

In any inelastic collision, the states of the colliding objects change, and the sum of their internal energies increases by an amount equal to the decrease in the sum of their kinetic energies.

Whenever some kinetic energy disappears in a collision, we can *always* determine changes in state to account for that loss. The big appeal of the relationship between state and internal energy is that we can now say:

The energy of a system of two colliding objects does not change during the collision.

This statement holds for all types of collisions: elastic, inelastic, and totally inelastic. In the elastic collision shown in **Figure 5.5a**, the collision alters the velocities and thus the kinetic energies of both carts. However, the sum of the kinetic energies before and after the collision remains unchanged, as shown by the kinetic energy bars. Because the collision is elastic, there are no changes in the states of the carts, which means their internal energies also remain unchanged, as shown by the internal energy bars. So the energy of the two-cart system does not change.

In the totally inelastic collision of Figure 5.5*b*, the sum of the kinetic energies decreases. Because changes occur in the states—the cars change shape—the internal energies of the carts change as well. The changes in motion and state are such that the energy of the system is the same before and after the collision. We account for the lost kinetic energy by saying that it has been converted to internal energy.

Note that in making the bar diagrams in Figure 5.5, I had to choose some initial value for the internal energies of the colliding objects. As we have no way (yet) of calculating internal energies, the values I chose are arbitrary. At present we are interested only in *changes* in energy, however, and so the initial value is not important.

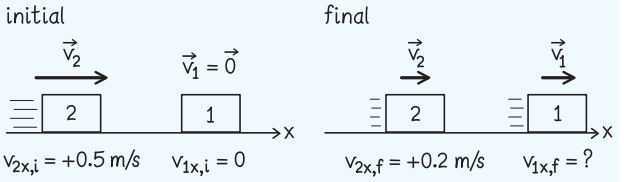
5.5 A piece of dough is thrown at a wall and sticks to it. Does the internal energy of the dough-wall system increase, decrease, or stay the same?

EXAMPLE 5.3 Internal energy change

A 0.2-kg cart 1 initially at rest is struck by an identical cart 2 traveling at $v_{2x,i} = +0.5$ m/s along a low-friction track. After the collision, the velocity of cart 2 is reduced to $v_{2x,f} = +0.2$ m/s. (*a*) Is the collision elastic, inelastic, or totally inelastic? (*b*) By what amount does the internal energy of the two-cart system change? (*c*) Make a bar diagram showing the initial and final kinetic and internal energies of the two carts.

1 GETTING STARTED I begin by organizing the information given in the problem in a sketch (**Figure 5.6**). To classify the collision, I need to determine the final relative speed, but the final velocity of cart 1 is not given.

Figure 5.6



2 DEVISE PLAN The two-cart system is isolated, and so the momentum of the system does not change, regardless of the type of collision. I can use this information to determine the final velocity of cart 1 and the final relative speed of the carts. By comparing the final and initial relative speeds, I can determine the type of collision. Once I know the initial and final velocities, I can calculate the kinetic energies using $K = \frac{1}{2}mv^2$ and determine what fraction of the initial kinetic energy has been converted to internal energy.

CONCEPTS

3 EXECUTE PLAN (*a*) The initial relative speed is

$$|v_{2x,i} - v_{1x,i}| = |(+0.5\text{ m/s}) - 0| = 0.5\text{ m/s}.$$

To determine $v_{1x,f}$, I apply conservation of momentum to the system. The initial momentum of the system is

$$(0.2\text{ kg})(+0.5\text{ m/s}) + (0.2\text{ kg})(0) = (0.2\text{ kg})(+0.5\text{m/s})$$

and its final momentum is

$$(0.2\text{ kg})(+0.2\text{ m/s}) + (0.2\text{ kg})(v_{1x,f}).$$

Conservation of momentum requires these two momenta to be equal:

$$(0.2\text{kg})(+0.5\text{m/s}) = (0.2\text{kg})(+0.2\text{m/s}) + (0.2\text{kg})(v_{1x,f})$$

$$(+0.5\text{ m/s}) = (+0.2\text{ m/s}) + v_{1x,f}$$

$$v_{1x,f} = +0.3\text{ m/s}.$$

The final relative speed is thus

$$|v_{2x,f} - v_{1x,f}| = |(+0.2\text{ m/s}) - (+0.3\text{ m/s})| = 0.1\text{ m/s},$$

which is different from the initial value. Thus the collision is inelastic. (I know that the collision is not *totally* inelastic because the relative speed has not been reduced to zero.) ✔

(*b*) The initial kinetic energies are

$$K_{1i} = 0$$

$$K_{2i} = \tfrac{1}{2}(0.2\text{ kg})(0.5\text{ m/s})^2 = 0.025\text{ kg}\cdot\text{m}^2/\text{s}^2$$

so $K_i = K_{1i} + K_{2i} = 0.025\text{ kg}\cdot\text{m}^2/\text{s}^2$.

The final kinetic energies are

$$K_{1f} = \tfrac{1}{2}(0.2\text{ kg})(0.3\text{ m/s})^2 = 0.009\text{ kg}\cdot\text{m}^2/\text{s}^2$$

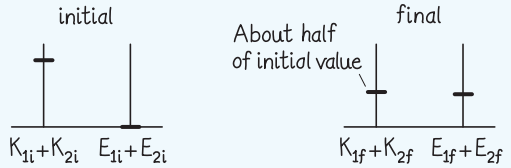
$$K_{2f} = \tfrac{1}{2}(0.2\text{ kg})(0.2\text{ m/s})^2 = 0.004\text{ kg}\cdot\text{m}^2/\text{s}^2$$

so $K_f = K_{1f} + K_{2f} = 0.013\text{ kg}\cdot\text{m}^2/\text{s}^2$.

The kinetic energy of the system has changed by an amount $(0.013\text{ kg}\cdot\text{m}^2/\text{s}^2) - (0.025\text{ kg}\cdot\text{m}^2/\text{s}^2) = -0.012\text{ kg}\cdot\text{m}^2/\text{s}^2$. To keep the energy of the system (the sum of its kinetic and internal energies) unchanged, the decrease in kinetic energy must be made up by an increase in internal energy. This tells me that the internal energy of the system increases by $0.012\text{ kg}\cdot\text{m}^2/\text{s}^2$. ✔

(*c*) My bar diagram is shown in **Figure 5.7**. The final kinetic energy bar is about half of the initial kinetic energy bar. Because I don't know the value of the initial internal energy, I set it to zero and make the final internal energy bar equal in height to the difference in the kinetic energy bars. ✔

Figure 5.7

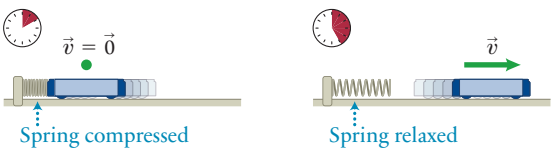


4 EVALUATE RESULT If the collision were elastic, the velocities of the carts would be interchanged and cart 2 would come to a stop (see Figure 4.5). So the collision must be inelastic. Indeed, I found that both the relative speed and the sum of the kinetic energies change in the collision, as expected for an inelastic collision.

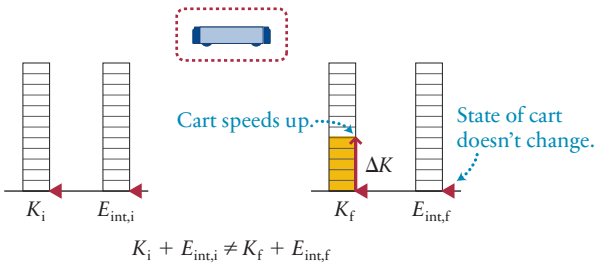
We are now in a position to extend the idea of internal energy to other interactions. Consider, for example, a cart initially at rest on a low-friction track set in motion by an expanding spring that is held fixed at one end, as shown in **Figure 5.8*a***. As the cart is accelerated by the spring, the cart's kinetic energy increases but its state doesn't change, and so its energy increases (Figure 5.8*b*). Where did this energy come from? The spring sets the cart in motion, and so it makes sense to assume that the spring transfers energy to the cart. Indeed, the spring expands—its state changes—and so its internal energy changes. If we include the spring in the system and attribute the increase in the cart's kinetic energy to a decrease in the spring's internal energy, we can again arrange things in such a way that the energy of the cart-spring system remains unchanged (Figure 5.8*c*).

Figure 5.8 Initial and final energies for two choices of system.

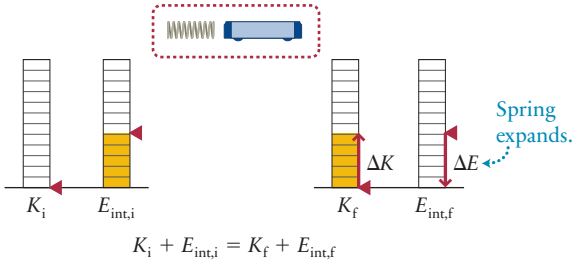
(*a*) Expanding spring accelerates cart from rest



(*b*) Initial and final energies: system = cart only



(*c*) Initial and final energies: system = cart + spring



CONCEPTS

If we replace the cart in Figure 5.8 with a cart with a different inertia, compress the spring to the same initial state, and then let that cart go, we discover that its final kinetic energy is exactly the same as that of the first cart. So the spring always transfers the same amount of energy as it expands from a given compressed state to its relaxed state *regardless* of the object to which it transfers that energy.

5.6 Think of a few other ways to put the cart of Figure 5.8 in motion. In each case, can you account for the increase in the cart's kinetic energy by either a change in state or a change in motion of another object?

Checkpoint 5.6 suggests that a change in the cart's kinetic energy can always be attributed to either a change in state (and therefore a change in the internal energy) or a change in the motion of another object. Because the states of objects can change in many different ways, we associate different forms of internal energy with different kinds of state change. For example, the internal energy associated with a change in chemical state is called *chemical energy* and the internal energy associated with a change in the temperature of an object is called *thermal energy*. **Table 5.3** lists additional forms of internal energy.

Absolutely everything happening around us entails changes in state and therefore changes in internal energy. Rivers flowing, air masses moving, machines lifting things, people walking, and atoms emitting light can all be expressed in terms of changes in state (and therefore changes in internal energy). More important, we discover that any change in state or motion is always accompanied by a compensating change in state or motion, and we can always attribute energy to these changes in such a way that the energy of the system remains unchanged. In other words, energy cannot be destroyed or created, and energy, like momentum, is a conserved quantity. Indeed, no observation has ever been found to violate the law of **conservation of energy:**

> **Energy can be transferred from one object to another or converted from one form to another, but it cannot be destroyed or created.**

Table 5.3 Various state changes and their associated internal energy

State change	Internal energy
temperature change	thermal energy
chemical change	chemical energy
reversible change in shape	elastic energy
phase transformation	transformation energy

Kinetic energy and all forms of internal energy are thus different manifestations of the same conserved quantity: *energy*.

5.7 (*a*) Is the momentum of the cart-spring system in Figure 5.8 constant? (*b*) Is the system isolated?

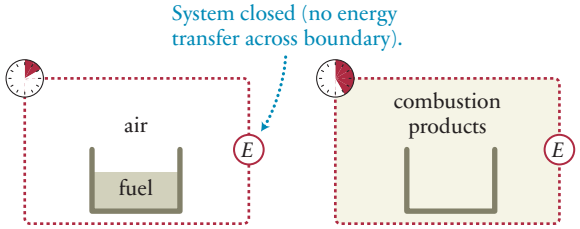
5.4 Closed systems

As Checkpoint 5.7 shows, the system comprising the spring and cart in Figure 5.8 is not isolated. However, no energy is transferred to it, and therefore the energy of the system is constant.* Any system to or from which no energy is transferred is called a **closed system.** An important point to keep in mind is that a closed system need not be isolated (and likewise an isolated system is not necessarily closed). The procedure for choosing a closed system is described in the Procedure box. To see how this procedure works, let's look at the setup in **Figure 5.9**—some fuel (gasoline, say) being burned in a container open to the air.

I begin by making a sketch of the canister with the fuel before and after the combustion (Figure 5.9*a*). The combustion involves two changes in state: a change in the chemical state of the fuel and the air, and a change

Figure 5.9 (*a*) The burning of fuel involves a change in chemical composition and in temperature. (*b*) If we choose a closed system, the change in chemical energy must be compensated by a change in thermal energy.

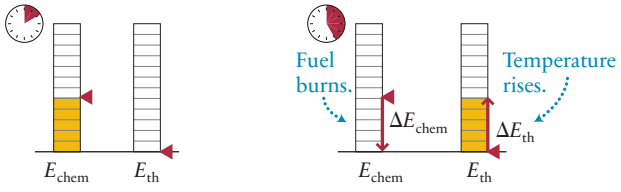
(*a*) Sketch initial and final conditions, identify changes, choose system



change in motion: none

changes in state:
- chemical state of fuel and air changes
- temperature rises

(*b*) Draw energy bar diagrams for initial and final conditions



*How do I know that no energy is transferred to the cart-spring system? The expanding spring and the accelerating cart do not cause any changes in the state or motion of the environment (the track, Earth, and so on). Consequently the energy of the environment doesn't change, which means that no energy has been transferred from the environment to the system.

CONCEPTS

PROCEDURE: Choosing a closed system

When we analyze energy changes, it is convenient to choose a system for which no energy is transferred to or from the system (a closed system). To do so, follow this procedure:

1. Make a sketch showing the initial and final conditions of the objects under consideration.
2. Identify all the changes in state or motion that occur during the time interval of interest.
3. Choose a system that includes all the objects undergoing these changes in state or motion. Draw a dashed line around the objects in your chosen system to represent the system boundary. Write "closed" near the system boundary to remind yourself that no energy is transferred to or from the system.
4. Verify that nothing in the surroundings of the system undergoes a change in motion or state that is related to what happens inside the system.

Once you have selected a closed system, you know that its energy remains constant.

in the temperature. The fuel and the air immediately surrounding it are what undergo these changes in state, and so I include them in my system. Because there are no other related changes in state, I know that my system is closed and its energy remains constant.

I can now make a bar diagram to represent the changes in energy that take place inside the system (Figure 5.9*b*). The change in the chemical state corresponds to a change in chemical energy, and the change in the temperature corresponds to a change in thermal energy. I therefore draw two bars. The combustion raises the temperature, and so the thermal energy increases. Because the system is closed and its energy remains constant, there must be a decrease in chemical energy that compensates for the increase in thermal energy.

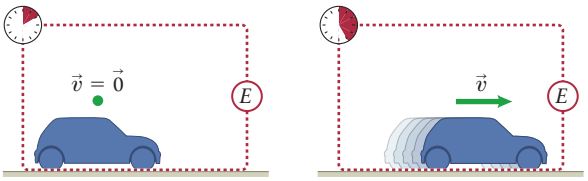
Given that the energy of a closed system remains constant, we can focus on the energy conversions and transfers that happen inside the system. We speak of an *energy conversion* when energy is converted from one form to another. In the combustion illustrated in Figure 5.9, for example, an amount ΔE_{chem} of chemical energy is converted to an equal amount ΔE_{th} of thermal energy. When energy is transferred from one object to another, we speak of an *energy transfer*. An example is shown in Figure 5.3*a*, where energy is transferred from cart 2 to cart 1.

Because we are still lacking the quantitative tools, let me illustrate the transfer and conversion of energy qualitatively with a few examples. By converting some form of internal energy to kinetic energy, we can put objects in motion. For example, a car burns gasoline as it accelerates from rest. If we ignore the effects of friction and air resistance, this situation represents a change in the chemical state of the gasoline and a change in the motion of the car. Chemical energy (a form of internal energy associated with the chemical state) stored in the gasoline is converted to kinetic energy of the car (**Figure 5.10**).

5.8 In describing what's going on in Figure 5.10, I ignored the change in temperature of the engine that accompanies the combustion of the fuel. Make a bar diagram that includes this change in temperature.

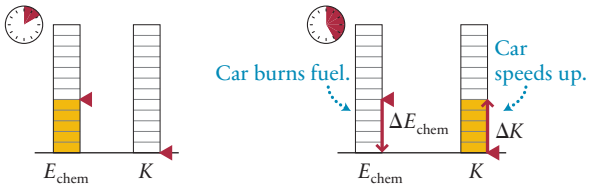
Figure 5.10 (*a*) An accelerating car constitutes a closed system because no changes occur in its environment. (*b*) The increase in the car's kinetic energy can be attributed to a decrease in chemical energy (due to the combustion of fuel).

(*a*) Initial and final conditions, changes in state and motion, system



change in motion: car accelerates
change in state: chemical state of fuel changes

(*b*) Energy bar diagrams for initial and final conditions



The generation and consumption of electrical power provide many other examples of energy conversions and transfers. In a coal-burning electrical power plant, coal is burned in a boiler to heat water and convert it to steam. The two complementary changes in state (a change in the chemical state of the coal and a change in the temperature and state of the water) correspond to a conversion of chemical energy to thermal energy. The steam drives a turbine, where the steam cools and a turbine blade spins, and so thermal energy is converted to kinetic energy. The moving turbine drives a generator, which converts the kinetic energy to electrical energy. Power lines transfer energy from the plant to our homes, where it is converted to many forms. Lamps convert electrical energy to light (a form of energy) and thermal energy (lamps get hot), a stereo converts electrical energy to sound energy, and on and on.

CONCEPTS

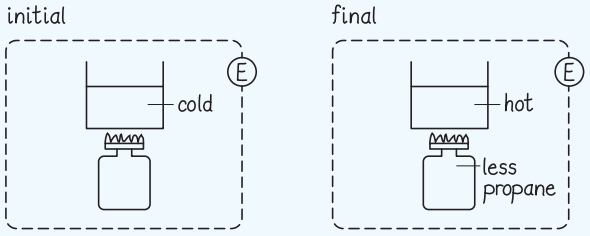
EXERCISE 5.4 State changes and internal energy

Choose an appropriate closed system and make a bar diagram representing the energy conversions and transfers that occur when (*a*) a pan of water is heated on a propane burner, (*b*) a cyclist accelerates from rest, and (*c*) a spring-loaded gun fires a ball of putty.

SOLUTION For each case, I apply the steps of the Procedure box on page 157. My sketches are shown in the figures.

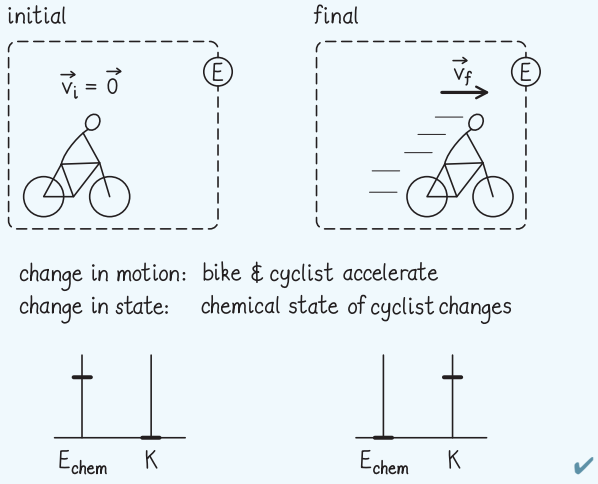
(*a*) Changes in motion: none. Changes in state: the temperature of the water increases, and the chemical states of the propane and the air change. The bar diagram shows an increase in thermal energy and an equal decrease in chemical energy (**Figure 5.11**). Chemical energy is converted to thermal energy, and in the process energy is transferred from the propane to the water.

Figure 5.11



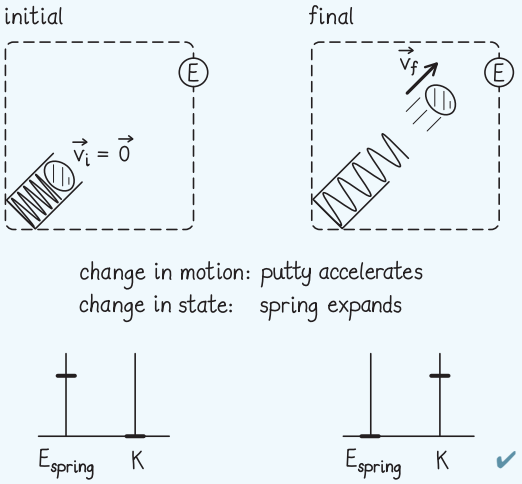
(*b*) Changes in motion: the bicycle and the cyclist accelerate. Changes in state: the chemical state of the cyclist changes because setting the bicycle in motion requires muscles to contract, a physiological process that involves a complex series of chemical reactions. The bar diagram shows an increase in kinetic energy and an equal decrease in chemical energy (**Figure 5.12**). Chemical energy is converted to kinetic energy of the bicycle and cyclist.

Figure 5.12



(*c*) Changes in motion: the putty is accelerated. Changes in state: the spring expands. The bar diagram shows a decrease in the elastic energy of the spring and an increase in the kinetic energy of the putty (**Figure 5.13**). As the spring expands, elastic energy is converted to kinetic energy of the putty. ✔

Figure 5.13



We started our study of physics by describing motion in terms of velocities, accelerations, and inertias. Now we have two new quantities, *momentum* and *energy*, and we have found two fundamental conservation laws that allow us to use simple accounting principles to describe changes in motion.

5.9 (*a*) Can the magnitude of the momentum of an object change without a change in the object's kinetic energy? (*b*) Without a change in the object's energy? (*c*) What are your answers to parts *a* and *b* for a system consisting of more than one object?

CONCEPTS

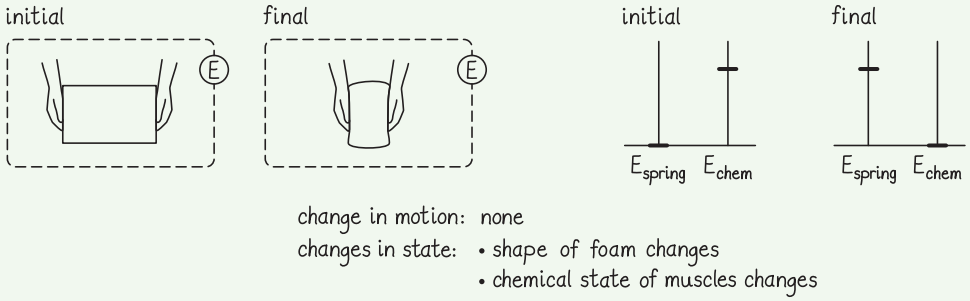
Self-quiz

1. Consider an isolated object at rest in space. The object contains internal energy in some form or another. Is it—in principle—possible to convert the internal energy to kinetic energy so that the object starts to move?
2. Imagine squeezing a piece of foam with your hands. Choose an appropriate closed system and make a bar diagram representing the energy conversions and transfers that occur during the squeezing.
3. When you heat a pot of water on a gas stove, the water temperature increases until the water begins to boil. This change in thermal state from cool water to hot water is due to chemical energy from the burning gas being converted to thermal energy of the water. Once boiling starts, the water temperature stays constant until all the water has turned to steam, even though the burning gas continues to transfer energy to the water. While the water is boiling off and becoming steam, what becomes of the energy released by the burning gas?
4. An electric fan turns electrical energy into wind energy (a form of kinetic energy because it involves moving air). Suppose a blowing fan is suddenly unplugged. Even though the fan no longer receives electrical energy, it continues to blow air while the blades slowly come to a stop. What type of energy is converted to wind energy after the fan is unplugged?

Answers

1. No. Getting the object to move would violate the law of conservation of momentum because the object would start with zero momentum ($\vec{p} = m\vec{v} = m\vec{0} = \vec{0}$) and end with nonzero momentum ($\vec{p} = m\vec{v}$).
2. Draw the foam before and after the squeezing. Two changes in state occur: The shape of the foam changes, and the chemical state of your muscles changes. These two changes in state correspond to changes in the foam's elastic energy and in chemical energy. In your closed system, include the foam and yourself (**Figure 5.14**).

Figure 5.14



3. It goes into changing the phase of the water from liquid to steam.
4. The kinetic energy of the fan blades is converted to wind energy.

5.5 Elastic collisions

As we saw in Section 5.1, we can classify collisions according to what happens to the relative velocity of the two colliding objects. The **relative velocity** of cart 2 relative to cart 1 is defined as

$$\vec{v}_{12} \equiv \vec{v}_2 - \vec{v}_1. \tag{5.1}$$

The relative velocity of cart 1 relative to cart 2 is the negative of this vector:

$$\vec{v}_{21} \equiv \vec{v}_1 - \vec{v}_2 = -\vec{v}_{12}. \tag{5.2}$$

If the magnitude of this relative velocity (the objects' **relative speed** $v_{12} \equiv |\vec{v}_2 - \vec{v}_1|$) is the same before and after the collision, the collision is elastic:

$$v_{12i} = v_{12f} \quad \text{(elastic collision)}. \tag{5.3}$$

For two objects moving along the x axis, we can write this as

$$v_{2x,i} - v_{1x,i} = -(v_{2x,f} - v_{1x,f}) \quad \text{(elastic collision)}, \tag{5.4}$$

where I have added a minus sign on the right because if object 2 initially moves faster than object 1, object 1 must move faster after the collision (**Figure 5.15**). In other words, in an elastic collision, if $v_{2x,i} > v_{1x,i}$, we must have $v_{2x,f} < v_{1x,f}$. The relative velocity in an elastic collision always changes sign after the collision: $v_{12x,i} = -v_{12x,f}$.

Because we can consider a system made up of two colliding objects to be isolated during the collision, conservation of momentum requires that the momentum of the colliding objects remains unchanged in the collision:

$$m_1 v_{1x,i} + m_2 v_{2x,i} = m_1 v_{1x,f} + m_2 v_{2x,f} \quad \text{(isolated system)}. \tag{5.5}$$

Equations 5.4 and 5.5 allow us to determine the final velocities of the objects, given their initial velocities.

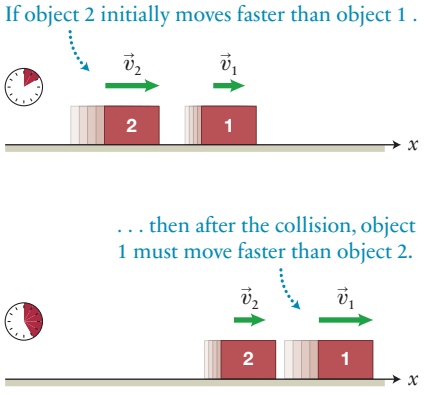


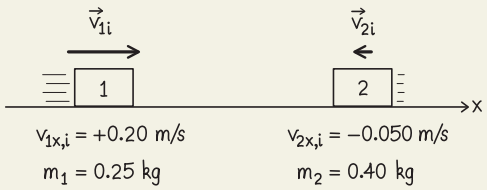
Figure 5.15 A collision between two objects moving in the same direction at different speeds.

EXAMPLE 5.5 Elastic collision

Two carts, one of inertia $m_1 = 0.25$ kg and the other of inertia $m_2 = 0.40$ kg, travel along a straight horizontal track with velocities $v_{1x,i} = +0.20$ m/s and $v_{2x,i} = -0.050$ m/s. What are the carts' velocities after they collide elastically?

❶ **GETTING STARTED** I begin by organizing the information in a sketch to help visualize the situation (**Figure 5.16**). I need to determine the velocities after the collision from the information given, knowing that the collision is elastic.

Figure 5.16



❷ **DEVISE PLAN** To determine the two unknowns $v_{1x,f}$ and $v_{2x,f}$, I need two equations. Because the collision is elastic, I can use Eqs. 5.4 and 5.5.

❸ **EXECUTE PLAN** I begin by solving Eq. 5.4 for $v_{2x,f}$:

$$v_{2x,f} = v_{1x,i} - v_{2x,i} + v_{1x,f}$$

so that I can eliminate $v_{2x,f}$ from Eq. 5.5:

$$m_1 v_{1x,i} + m_2 v_{2x,i} = m_1 v_{1x,f} + m_2(v_{1x,i} - v_{2x,i} + v_{1x,f}).$$

Solving this expression for $v_{1x,f}$ yields

$$v_{1x,f} = \frac{m_1 - m_2}{m_1 + m_2} v_{1x,i} + \frac{2m_2}{m_1 + m_2} v_{2x,i}. \tag{1}$$

QUANTITATIVE TOOLS

Repeating this procedure, I can eliminate $v_{1x,f}$ and solve for $v_{2x,f}$:

$$v_{2x,f} = \frac{2m_1}{m_1 + m_2} v_{1x,i} - \frac{m_1 - m_2}{m_1 + m_2} v_{2x,i}. \quad (2)$$

Substituting the given values into Eqs. 1 and 2 yields $v_{1x,f} = -0.11$ m/s and $v_{2x,f} = +0.14$ m/s. ✔

4 EVALUATE RESULT The final velocities yield the final relative speed $v_{12f} = |(+0.14\,\text{m/s}) - (-0.11\,\text{m/s})| = 0.25\,\text{m/s}$. This value is the same as the initial relative speed $v_{12i} = |(-0.05\,\text{m/s}) - (+0.20\,\text{m/s})| = 0.25\,\text{m/s}$, as required for an elastic collision.

Let me now show how Eqs. 5.4 and 5.5 lead to an expression that shows that the kinetic energy of two colliding objects doesn't change in an elastic collision. Rearranging the terms in Eq. 5.5, I get

$$m_1(v_{1x,i} - v_{1x,f}) = m_2(v_{2x,f} - v_{2x,i}), \quad (5.6)$$

while Eq. 5.4 yields

$$v_{1x,i} + v_{1x,f} = v_{2x,i} + v_{2x,f}. \quad (5.7)$$

Now I multiply the left side of Eq. 5.6 by the left side of Eq. 5.7 and likewise for the right sides to get

$$m_1(v_{1x,i} - v_{1x,f})(v_{1x,i} + v_{1x,f}) = m_2(v_{2x,f} - v_{2x,i})(v_{2x,i} + v_{2x,f}), \quad (5.8)$$

which multiplies out to

$$m_1 v_{1x,i}^2 - m_1 v_{1x,f}^2 = m_2 v_{2x,f}^2 - m_2 v_{2x,i}^2. \quad (5.9)$$

For motion along the x axis, the square of the x component of a velocity is the same as the square of the speed, $v_x^2 = v^2$, and so I can drop all the subscripts x. If I put all the v_i terms on the left and all the v_f terms on the right, I get

$$m_1 v_{1i}^2 + m_2 v_{2i}^2 = m_1 v_{1f}^2 + m_2 v_{2f}^2. \quad (5.10)$$

I now divide both sides by 2 to get

$$\tfrac{1}{2} m_1 v_{1i}^2 + \tfrac{1}{2} m_2 v_{2i}^2 = \tfrac{1}{2} m_1 v_{1f}^2 + \tfrac{1}{2} m_2 v_{2f}^2. \quad (5.11)$$

As you saw in Section 5.2, the quantity $\frac{1}{2}mv^2$ is the **kinetic energy**, which we denote by K:*

$$K \equiv \tfrac{1}{2} m v^2. \quad (5.12)$$

With this definition, Eq. 5.11 simplifies to

$$K_{1i} + K_{2i} = K_{1f} + K_{2f} \quad (\text{elastic collision}), \quad (5.13)$$

where $K_{1i} = \frac{1}{2} m_1 v_{1i}^2$ is the initial kinetic energy of object 1, and so on. In words, in an elastic collision the sum of the initial kinetic energies is the same as the sum of the final kinetic energies.

*The factor $\frac{1}{2}$ in Eq. 5.12 is there for convenience. We could have defined kinetic energy as simply mv^2, but the definition given in Eq. 5.12 simplifies matters later on.

Note that there is no new physics in Eqs. 5.11 and 5.13—both express the same fact as Eq. 5.3. The great advantage of writing this information in the form of Eq. 5.13 is that this form expresses Eq. 5.3 in terms of an extensive quantity. Because kinetic energy is an extensive quantity, the kinetic energy of a system comprising two objects is the sum of the kinetic energies of the individual objects: $K = K_1 + K_2$. Equation 5.13 thus tells us that the kinetic energy K of a system of objects undergoing an elastic collision does not change:

$$K_i = K_f \quad (\text{elastic collision}). \tag{5.14}$$

This result can also be written in the form

$$\Delta K = 0 \quad (\text{elastic collision}). \tag{5.15}$$

As you know from studying Table 5.1, the SI unit for kinetic energy is $\text{kg}\cdot\text{m}^2/s^2$. Because energy is such an important quantity, we give this combination of units the name **joule** (rhymes with *pool*), after the British physicist James Prescott Joule (1818–1889), who studied energy:

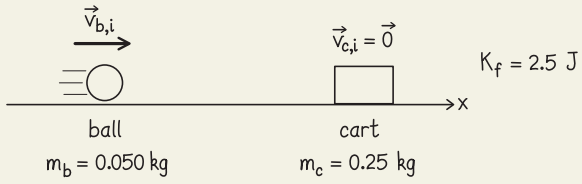
$$1\ \text{kg}\cdot\text{m}^2/\text{s}^2 \equiv 1\ \text{J}.$$

EXAMPLE 5.6 Collision and kinetic energy

A rubber ball of inertia $m_b = 0.050$ kg is fired along a track toward a stationary cart of inertia $m_c = 0.25$ kg. The kinetic energy of the system after the two collide elastically is 2.5 J. (*a*) What is the initial velocity of the ball? (*b*) What are the final velocities of the ball and the cart?

❶ **GETTING STARTED** I organize the problem graphically (**Figure 5.17**). I choose my x axis in the direction of the incoming rubber ball. Only one initial velocity is given. I need to determine the other initial velocity and both final velocities.

Figure 5.17



❷ **DEVISE PLAN** Because the collision is elastic, I know that the kinetic energy of the system doesn't change, which means that the final value (2.5 J) is the same as the initial value. Because the cart is initially at rest, all of this kinetic energy belongs initially to the ball, and so I can use Eq. 5.12 to determine the initial velocity of the ball. Once I have this information, I know the initial velocities of both colliding objects and I can use Eqs. 1 and 2 from Example 5.5 to calculate the final velocities.

❸ **EXECUTE PLAN** (*a*) From Eq. 5.12, I obtain the initial speed of the ball:

$$v_{b,i} = \sqrt{\frac{2K_{b,i}}{m_b}} = \sqrt{\frac{2(2.5\ \text{J})}{0.05\ \text{kg}}} = 10\ \text{m/s}.$$

Because the ball is initially moving in the positive x direction, its initial velocity is given by $v_{bx,i} = +10$ m/s. ✔

(*b*) I can now substitute the two initial velocities and the inertias into Eqs. 1 and 2 of Example 5.5, which gives me $v_{bx,f} = -6.7$ m/s and $v_{cx,f} = +3.3$ m/s. ✔

❹ **EVALUATE RESULT** It makes sense that the velocity of the ball is reversed by the collision because the inertia of the cart is so much greater than that of the ball. Now that I know both the initial and final velocities, I can also check to make sure that the relative speed remains the same. Because the cart is initially at rest, the initial relative speed is 10 m/s; the final relative speed is $v_{bc,f} = |(+3.3\ \text{m/s}) - (-6.7\ \text{m/s})| = 10$ m/s, which is the same, as required for an elastic collision.

5.10 Conservation of momentum tells us that mv_x doesn't change for an isolated system. Now we see that in an elastic collision mv_x^2 also doesn't change. Does this mean that mv_x^3 remains unchanged, too?

QUANTITATIVE TOOLS

5.6 Inelastic collisions

In a totally inelastic collision between two objects, the objects move together after the collision and so the two final velocities are identical. This means that the final relative speed is zero:

$$v_{12f} = 0 \quad (\text{totally inelastic collision}). \tag{5.16}$$

Therefore if you know the inertias m_1 and m_2 of the objects and the x components of their initial velocities $v_{1x,i}$ and $v_{2x,i}$, you can obtain the final velocities from Eq. 5.5 by setting

$$v_{1x,f} = v_{2x,f} \quad (\text{totally inelastic collision}). \tag{5.17}$$

The majority of collisions, however, fall somewhere between the two extremes of elastic and totally inelastic. For these collisions, the final relative speed is between zero and the value of the initial relative speed. For these cases, it is convenient to define the ratio of relative speeds as

$$e \equiv \frac{v_{12f}}{v_{12i}}. \tag{5.18}$$

The unitless quantity e is called the **coefficient of restitution** of the collision. It tells us how much of the initial relative speed is restituted (restored) after the collision. Because it is a ratio of speeds (which are always positive), e is always positive or zero.

In general, it is more convenient to write Eq. 5.18 in terms of components:

$$e = -\frac{v_{2x,f} - v_{1x,f}}{v_{2x,i} - v_{1x,i}} = -\frac{v_{12x,f}}{v_{12x,i}}, \tag{5.19}$$

where the minus sign appears because the relative velocity changes sign after the collision (see Figure 5.15).

When $e = 1$, Eq. 5.19 is equal to Eq. 5.4 and the collision is elastic (all of the initial relative speed is restored). A value of $e = 0$ means that v_{12f} must be zero, and so the collision is totally inelastic. For values of e between 0 and 1, the collision is inelastic. This information is summarized in Table 5.4.

In an inelastic collision, therefore, we again have two equations—Eqs. 5.1 and 5.19—but we need to know the value of e to be able to calculate the final velocities. The details of how the states of the objects change in the collision (these state changes depend on the material and structural properties of the objects) are hidden in this coefficient e.

Table 5.4 Coefficient of restitution for various processes

Process	Relative speed	Coefficient of restitution
totally inelastic collision	$v_{12f} = 0$	$e = 0$
inelastic collision	$0 < v_{12f} < v_{12i}$	$0 < e < 1$
elastic collision	$v_{12f} = v_{12i}$	$e = 1$
explosive separation*	$v_{12f} > v_{12i}$	$e > 1$

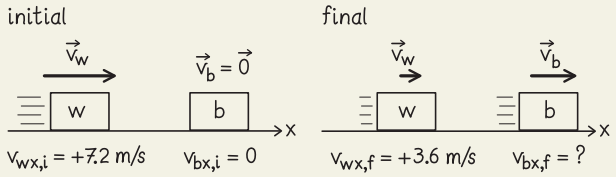
*See Section 5.8.

EXAMPLE 5.7 Restitution

A white car of inertia 1200 kg that is moving at a speed of 7.2 m/s rear-ends a blue car of inertia 1000 kg that is initially at rest. Immediately after the collision, the white car has a speed of 3.6 m/s. What is the coefficient of restitution for this collision?

❶ **GETTING STARTED** By putting all the information in graphical form (**Figure 5.18**), I see that I do not have a value for the final velocity of the blue car, which I need in order to calculate the coefficient of restitution.

Figure 5.18



❷ **DEVISE PLAN** Because the system of the two colliding cars is isolated, I can use conservation of momentum (Eq. 5.5) to obtain the final velocity of the blue car, $v_{bx,f}$.

❸ **EXECUTE PLAN** I begin by writing Eq. 5.5 with subscripts appropriate to this particular collision:

$$m_w v_{wx,i} + m_b v_{bx,i} = m_w v_{wx,f} + m_b v_{bx,f}$$

or $$m_w v_{wx,i} + 0 = m_w v_{wx,f} + m_b v_{bx,f}.$$

Solving for $v_{bx,f}$ and substituting the values given, I get

$$v_{bx,f} = \frac{m_w}{m_b}(v_{wx,i} - v_{wx,f})$$

$$= \frac{1200\text{ kg}}{1000\text{ kg}}(7.2\text{ m/s} - 3.6\text{ m/s}) = 4.3\text{ m/s}.$$

Now I can substitute the initial and final velocities into Eq. 5.19 to obtain the coefficient of restitution:

$$e = -\frac{v_{wx,f} - v_{bx,f}}{v_{wx,i} - v_{bx,i}}$$

$$= -\frac{(+3.6\text{ m/s}) - (+4.3\text{ m/s})}{(7.2\text{ m/s}) - 0} = 0.097. \checkmark$$

❹ **EVALUATE RESULT** The coefficient of restitution is rather small, but that is to be expected because cars don't bounce like superballs. Indeed, the blue car moves just a little bit faster than the white car after the collision (4.3 m/s versus 3.6 m/s), as I expect.



5.11 In a totally inelastic collision between two objects in an isolated system with one of the objects initially at rest, is it possible to lose *all* of the system's initial kinetic energy?

5.7 Conservation of energy

Because no energy is transferred to or from a closed system, conservation of energy requires that any change in the kinetic energy of a closed system be compensated by an equal change in the internal energy so that the sum of the kinetic and internal energies of the system does not change:

$$K_i + E_{int,i} = K_f + E_{int,f} \quad \text{(closed system)}. \tag{5.20}$$

In contrast to Eq. 5.14, which holds for only an elastic collision, Eq. 5.20 holds for any closed system. If we write the sum of the kinetic and internal energies of an object or system as the **energy** of the object or system,

$$E \equiv K + E_{int}, \tag{5.21}$$

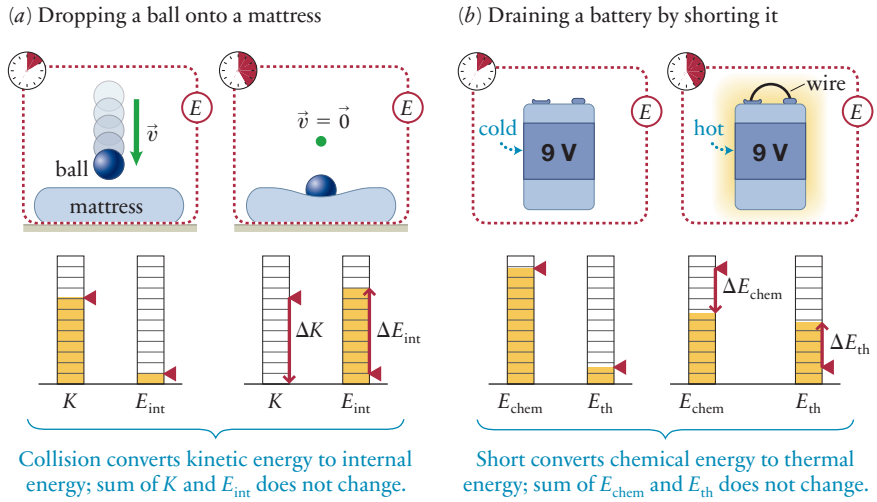
then we can rewrite Eq. 5.20 as

$$zE_i = E_f \quad \text{(closed system)} \tag{5.22}$$

or $$\Delta E = 0 \quad \text{(closed system)}. \tag{5.23}$$

Note the parallel between Eq. 4.17, $\Delta\vec{p} = \vec{0}$, and Eq. 5.23. Equation 4.17 embodies conservation of momentum and states that the momentum of an isolated system cannot change. Equation 5.23 embodies conservation

Figure 5.19 Two examples of energy conservation in a closed system. (*a*) The ball's kinetic energy is converted to internal energy in the deformed mattress. (*b*) Chemical energy stored in the battery is converted to thermal energy (the battery gets hot). In both cases, as the bar graphs show, the system's energy remains the same.



of energy: It states that the energy of a closed system cannot change; that is, energy cannot be created or destroyed.

Even though we cannot yet calculate the internal energy E_{int}, Eq. 5.23 allows us to draw a number of important conclusions. First, if the kinetic energy of a closed system changes, then the state of the system must change in such a way that its internal energy changes by an amount

$$\Delta E_{\text{int}} = -\Delta K \quad \text{(closed system)}. \tag{5.24}$$

As an example, consider **Figure 5.19*a***. A ball is dropped onto a mattress, where it comes to rest. During the time interval from the instant just before the ball hits the mattress to the instant at which the ball comes to rest, the motion of the ball changes and the shape of the mattress changes. The ball and mattress constitute a closed system, and so the decrease in kinetic energy must be accompanied by an increase in internal energy. Equation 5.24 requires the loss in kinetic energy to be equal to the gain in internal energy.

The second conclusion we can draw from Eq. 5.23 is that if the kinetic energy of a closed system does not change ($\Delta K = 0$), neither does the internal energy of the system ($\Delta E_{\text{int}} = 0$). Because E_{int} is the sum of the internal energies of all the parts that make up the system, however, $\Delta E_{\text{int}} = 0$ does *not* mean that no changes in state take place. For instance, E_{int} remains constant when internal energy is transferred from one part of a closed system to another:

$$\Delta E_{\text{int},1} = -\Delta E_{\text{int},2} \Rightarrow \Delta E_{\text{int}} = 0 \quad \text{(closed system, } \Delta K = 0\text{)} \tag{5.25}$$

or when one form of internal energy is converted to another:

$$\Delta E_{\text{form 1}} = -\Delta E_{\text{form2}} \Rightarrow \Delta E_{\text{int}} = 0 \quad \text{(closed system, } \Delta K = 0\text{)}. \tag{5.26}$$

As an example, consider Figure 5.19*b*. When the battery is drained rapidly, it becomes very hot (as does the wire). Because there is no motion before and after the draining, no change in kinetic energy occurs, but the chemical energy in the battery is converted to thermal energy. Equation 5.26 requires the loss in chemical energy to be equal to the gain in thermal energy:

$$\Delta E_{\text{chem}} + \Delta E_{\text{th}} = 0. \tag{5.27}$$

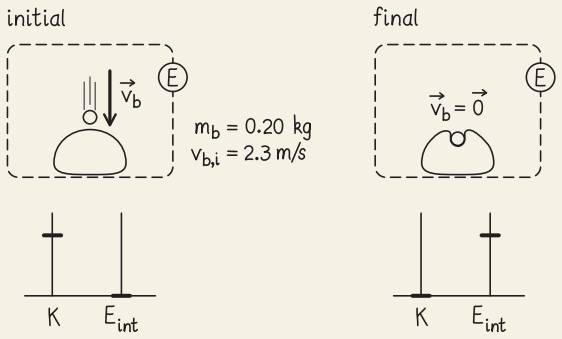
In practice there are often more than two simultaneous changes of state in a system (see Checkpoint 5.8, for instance), but regardless of how many energy conversions and transfers take place, the law of conservation of energy requires that the *amount of energy* in a closed system never changes.

EXAMPLE 5.8 Making light

A 0.20-kg steel ball is dropped into a ball of dough, striking the dough at a speed of 2.3 m/s and coming to rest inside the dough. If it were possible to turn all of the energy converted in this totally inelastic collision into light, how long could you light a desk lamp? It takes 25 J to light a desk lamp for 1.0 s.

❶ **GETTING STARTED** I begin by applying the procedure for choosing a closed system. Although the problem doesn't specify it explicitly, I'm assuming the dough is at rest both before and after the steel ball is dropped in it; it could, for example, be at rest on a countertop. Only the steel ball has kinetic energy initially, and all of this energy is converted to internal energy as the ball comes to rest in the dough (Figure 5.20). So I have to calculate the initial kinetic energy of the ball and determine how long that amount of energy could light a lamp, given that 25 J lights a lamp for 1.0 s.

Figure 5.20



❷ **DEVISE PLAN** To determine the initial kinetic energy of the ball, I use Eq. 5.12. Then I divide this result by 25 J to determine how many seconds I can light a lamp.

❸ **EXECUTE PLAN** The initial kinetic energy of the ball is

$$K_{b,i} = \tfrac{1}{2} m_b v_{b,i}^2 = \tfrac{1}{2}(0.20\text{ kg})(2.3\text{ m/s})^2 = 0.53\text{ J}.$$

Given that a desk lamp requires 25 J per second, this 0.53 J lights a lamp for

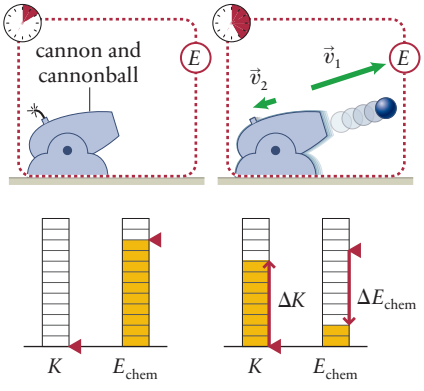
$$\frac{\text{energy available}}{\text{energy needed per second}} = \frac{0.53\text{ J}}{25\text{ J/s}} = 0.021\text{ s}. \checkmark$$

❹ **EVALUATE RESULT** The length of time I obtained, two hundreths of a second, is not very much! However, a 0.20-kg steel ball moving at 2.3 m/s does not have much kinetic energy: I know from experience that a small steel ball's ability to induce state changes—to crumple or deform objects, for example—is very limited. So it makes sense that one can't light a desk lamp for very long.

5.12 A gallon of gasoline contains approximately 1.2×10^8 J of energy. If all of this energy were converted to kinetic energy in a 1200-kg car, how fast would the car go?

5.8 Explosive separations

Figure 5.21 When a cannon is fired, internal (chemical) energy is converted to kinetic energy.

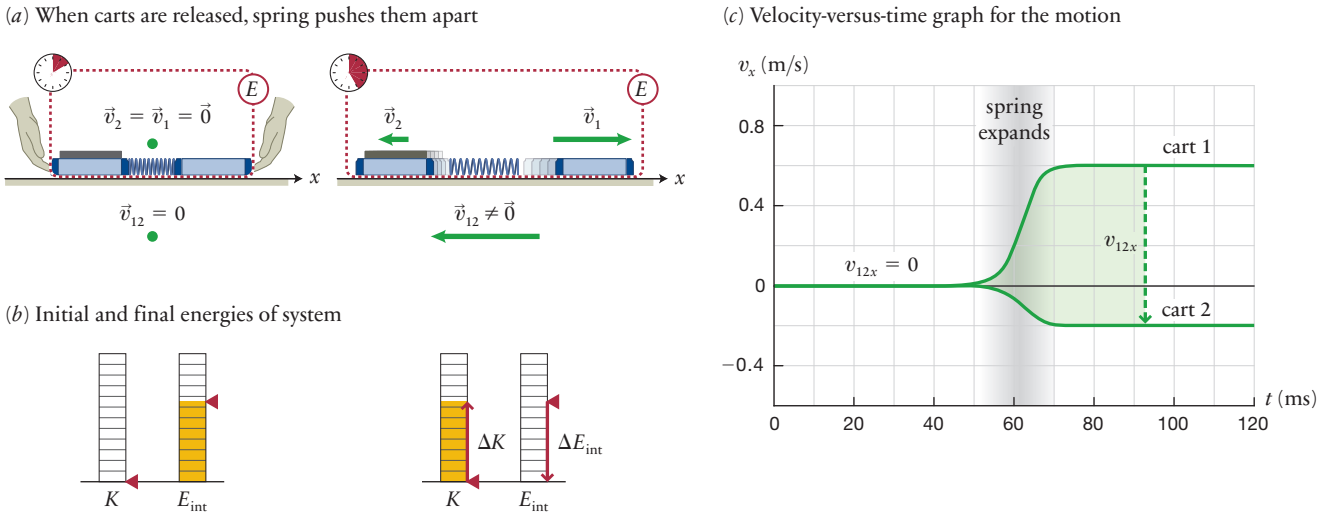


Is it possible to have a process in which kinetic energy is *gained* at the expense of internal energy? Yes, in an *explosion*, or any other type of **explosive separation**, where objects separate or break apart from each other, kinetic energy increases and internal energy decreases. The firing of a cannon is one example (Figure 5.21). Initially the cannon and cannonball are at rest. When the cannon is fired, the cannonball flies out of the barrel, and the cannon *recoils* in the opposite direction. (If it didn't recoil, the momentum of the system would not remain zero.) Both the cannon and cannonball therefore gain kinetic energy at the expense of chemical energy in the gun powder. This situation is the inverse of a totally inelastic collision: In this explosive separation, we start with the two objects together ($v_{12} = 0$) and end with them moving apart ($v_{12} > 0$).

Figure 5.22*a* shows an explosive separation that can be carried out on a low-friction track. Two carts, of inertias m and $3m$, are held against a compressed spring. When the carts are released, they move apart as the

QUANTITATIVE TOOLS

Figure 5.22 Another example of an explosive separation.



spring expands. As it expands, the spring's state changes and so does its internal energy; the decrease in internal energy of the spring causes an increase in the kinetic energy of the carts (Figure 5.22*b*). Notice how the velocity-versus-time graph for this explosive separation (Figure 5.22*c*) is the inverse of the one in Figure 5.2*b*.

To determine the final speeds in Figure 5.22, we need to know how much energy E_{int} the spring releases, a topic not covered until Chapter 9. Once we know ΔE_{int}, we have two equations that allow us to obtain the two final velocities, one a consequence of conservation of momentum:

$$0 = m_1 v_{1x,f} + m_2 v_{2x,f} \tag{5.28}$$

(because $v_{1x,i} = v_{2x,i} = 0$) and the other a consequence of conservation of energy:

$$\Delta K + \Delta E_{int} = \tfrac{1}{2} m_1 v_{1f}^2 + \tfrac{1}{2} m_2 v_{2f}^2 + \Delta E_{int} = 0. \tag{5.29}$$

Note that in this example the *initial* relative speed is zero, making e in Eq. 5.19 infinite. There is no restitution because there is no initial relative speed to be restituted. In the more general case where the two objects have a nonzero initial relative speed, $e > 1$.

EXAMPLE 5.9 Spring energy

A 0.25-kg cart is held at rest against a compressed spring as in Figure 5.8*a* and then released. The cart's speed after it separates from the spring is 2.5 m/s. The spring is then compressed by the same amount between a 0.25-kg cart and a 0.50-kg cart, as shown in Figure 5.22*a*, and the carts are released from rest. What are the carts' speeds after separating from the spring?

❶ **GETTING STARTED** The key point in this problem is the identical compression of the spring in the two cases: The initial state of the spring is therefore the same before both releases. Because the spring ends in the same uncompressed state in both cases, the change in its internal energy must be the same in both cases. In the first case, all of this energy is transferred to the 0.25-kg cart. In the second case, the same amount of energy is distributed between the two carts.

❷ **DEVISE PLAN** To calculate the kinetic energy of the single cart in the first release, I use Eq. 5.12. This gives me the amount of energy stored in the compressed spring. The final velocities of the two carts in the second case are then given by Eqs. 5.28 and 5.29.

❸ **EXECUTE PLAN** From Eq. 5.12, I get

$$K = \tfrac{1}{2} m v^2 = \tfrac{1}{2} (0.25\ \text{kg}) (2.5\ \text{m/s})^2 = 0.78\ \text{J}$$

and so the change in the spring's internal energy is $\Delta E_{int} = -0.78$ J. Next I rewrite Eq. 5.28 as $v_{1x,f} = -(m_2/m_1) v_{2x,f}$. Substituting this result in Eq. 5.29, I get

$$\tfrac{1}{2} m_1 \left(\frac{m_2}{m_1} \right)^2 v_{2x,f}^2 + \tfrac{1}{2} m_2 v_{2x,f}^2 = -\Delta E_{int}.$$

(Continued)

QUANTITATIVE TOOLS

Solving for the final velocity of cart 2 gives

$$v_{2xf} = \sqrt{\frac{-2m_1\Delta E_{int}}{m_2(m_1 + m_2)}}$$

$$v_{2xf} = \sqrt{\frac{-2(0.25\text{ kg})(-0.78\text{ J})}{(0.50\text{ kg})(0.25\text{ kg} + 0.50\text{ kg})}} = 1.0\text{ m/s.}\ \checkmark$$

Substituting this result into my rewritten Eq. 5.28, $v_{1x,f} = -(m_2/m_1)v_{2x,f}$, I get $v_{1x,f} = -2.0$ m/s. ✔

4 EVALUATE RESULT The carts move in opposite directions, as expected. I also note that cart 1 moves at twice the speed of cart 2, as it should to keep the final momentum of the system zero. Finally, because my assignment of m_1 and m_2 is arbitrary, I verify that I get the same result when I substitute $m_1 = 0.50$ kg and $m_2 = 0.25$ kg. (You may want to check this yourself. When you reverse the inertias, why does the velocity of cart 1 reverse to positive and the velocity of cart 2 reverse to negative?)



5.13 Does each cart in Example 5.9 get half of the spring's energy? Why or why not?

Chapter Summary

Kinetic energy (Sections 5.2, 5.5)

Concepts The **kinetic energy** of an object is the energy associated with its motion. Kinetic energy is a positive scalar quantity and is independent of the direction of motion.

Quantitative Tools The **kinetic energy** K of an object of inertia m moving at speed v is

$$K = \tfrac{1}{2}mv^2. \tag{5.12}$$

The SI unit of kinetic energy is the **joule** (J):

$$1\ \text{J} = 1\ \text{kg}\cdot\text{m}^2/\text{s}^2.$$

Relative velocity, states, and internal energy (Sections 5.1, 5.3, 5.5)

Concepts In a collision between two objects, the velocity of one object relative to the velocity of the other object is the **relative velocity** $\vec{v}_{12}$. The magnitude of the relative velocity is the **relative speed** v_{12}.

The **state** of an object is its condition as specified by some complete set of physical parameters. Energy associated with the object's state but not with its motion is called the **internal energy** of the object.

We can consider a system of two colliding objects to be isolated during the collision. Therefore the momentum of the system remains constant during all the collisions we study.

Quantitative Tools The **relative velocity** $\vec{v}_{12}$ of object 2 relative to object 1 is

$$\vec{v}_{12} \equiv \vec{v}_2 - \vec{v}_1. \tag{5.1}$$

The **relative speed** v_{12} of object 2 relative to object 1 is the magnitude of $\vec{v}_{12}$:

$$v_{12} = |\vec{v}_2 - \vec{v}_1|.$$

Because momentum is a conserved quantity, the momentum of a system remains constant during a collision:

$$p_{x,\text{i}} = p_{x,\text{f}}.$$

Types of collisions (Sections 5.5, 5.6, 5.8)

Concepts The **coefficient of restitution** e for a collision is a positive, unitless quantity that tells how much of the initial relative speed is restored after the collision.

For an **elastic collision,** the relative speed is the same before and after the collision, and the coefficient of restitution is equal to 1. The collision is **reversible,** and the kinetic energy of the system made up of the colliding objects is constant.

For an **inelastic collision,** the relative speed after the collision is less than it was before the collision. The coefficient of restitution is between 0 and 1, and the collision is **irreversible.** The kinetic energy of the objects changes during the collision, but the energy of the system does not change. If the objects stick together, the final relative speed is zero; the collision is **totally inelastic,** and the coefficient of restitution is 0.

Quantitative Tools The **coefficient of restitution** e is

$$e = \frac{v_{12\text{f}}}{v_{12\text{i}}} = -\frac{v_{2x,\text{f}} - v_{1x,\text{f}}}{v_{2x,\text{i}} - v_{1x,\text{i}}}. \tag{5.18, 5.19}$$

For an **elastic collision,**

$$v_{12\text{i}} = v_{12\text{f}} \tag{5.3}$$

$$K_\text{i} = K_\text{f} \tag{5.14}$$

$$e = 1.$$

For an **inelastic collision,**

$$v_{12\text{f}} < v_{12\text{i}}$$

$$K_\text{f} < K_\text{i}$$

$$0 < e < 1.$$

For a **totally inelastic collision,**

$$v_{12\text{f}} = 0 \tag{5.16}$$

$$e = 0.$$

For an **explosive separation,** kinetic energy is gained during the collision and the coefficient of restitution is greater than 1.

For an **explosive separation,**

$$v_{12f} > v_{12i}$$
$$K_f > K_i$$
$$e > 1.$$

Conservation of energy (Sections 5.3, 5.4, 5.7)

Concepts The energy of any system is the sum of the kinetic energies and internal energies of all the objects that make up the system.

The law of **conservation of energy** states that energy can be transferred from one object to another or converted from one form to another, but it cannot be destroyed or created.

A **closed system** is one in which no energy is transferred in or out. The energy of such a system remains constant.

Quantitative Tools The energy of a system is

$$E = K + E_{int}. \tag{5.21}$$

The law of conservation of energy requires the energy of a **closed system** to be constant:

$$E_i = E_f. \tag{5.22}$$

Questions and Problems

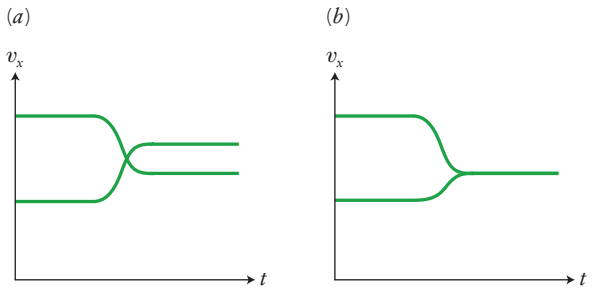
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

5.1 Classification of collisions

1. In a billiard game, two balls collide having inertia 0.25 kg each. Initially, ball 1 with velocity 8 m/s hits the ball 2 at rest. After collision, ball 1 moves with velocity 5 m/s and ball 2 moves with velocity 3 m/s. Comment whether the collision is elastic or inelastic. •
2. Suppose you have an isolated system in which two objects about to collide have equal and opposite momenta. If the collision is totally inelastic, what can you say about the motion after the collision? •
3. Figure P5.3 shows velocity-versus-time graphs for two situations in which a pair of objects collide. For each situation, decide whether the collision is elastic, inelastic, or totally inelastic. ••

Figure P5.3



4. When you drop a tennis ball on the floor from a height, it will bounce back. (*a*) Do you expect the ball to return to the same height? (*b*) Is this collision elastic or inelastic? Justify your answer. ••

5.2 Kinetic energy

5. The ratio of kinetic energies of two moving objects A and B is 1:5. Calculate the ratio of their momenta if the inertia of A is 5 times the inertia of B. •
6. As a result of an elastic collision between carts 1 and 2, the kinetic energy of cart 1 doubles. Does the kinetic energy of cart 2 change? If so, by what amount? •
7. A common energy unit used in food chemistry is the *food calorie* (1 Cal = 1000 cal = 4186 J). What must be the inertia of a person who is moving at a walking pace of 1.0 m/s and has kinetic energy numerically equal to the food energy in a jelly doughnut (300 Cal)? •
8. (*a*) Write an expression for the kinetic energy of an object in terms of its momentum p and inertia m. (*b*) If kinetic energy can be written in terms of momentum, how can the kinetic energy of a system change in an inelastic collision even though the law of momentum conservation forces the system's momentum to remain unchanged? ••
9. If two objects A and B have the same momentum but A has four times the kinetic energy of B, what is the ratio of their inertias? ••
10. The ratio of momenta of two objects A and B is 1: 4. The ratio of the kinetic energies of A and B is 1:2. Find the ratio of their velocities. ••
11. Object X has great inertia, and object Y has much less inertia. If the objects have the same momentum, which has more kinetic energy? If they have the same kinetic energy, which has greater momentum? ••

12. Draw a graph of kinetic energy versus distance fallen for a brick that is falling vertically from the top of a tall building. ••

5.3 Internal energy

13. You see a book on a table and give the book a push. It slides across the table and then comes to rest. Describe any changes in the physical state of the book-table system. •
14. What is the interaction type in each of the following cases: (*a*) two colliding marbles, (*b*) two cars in a collision in which they move together, (*c*) a bullet fired into and passing through a wooden block, (*d*) a bullet fired into a wooden block that gets embedded in it, and (*e*) one cart hitting another on a low-friction air track. •
15. What physical phenomenon is at work in the launching of a rocket? Is the performance of the rocket engine better on earth or in space? •
16. A small bullet is fired into a large piece of wood. After the bullet penetrates the wood, the assembly moves as one unit along a low-friction track in the direction of travel of the bullet. (*a*) After the bullet is stuck in the piece of wood, is the momentum of the wood (not including the bullet) *greater than, equal to,* or *less than* the initial momentum of the bullet? State any assumptions explicitly. (*b*) Is the combined kinetic energy of the bullet and the wood after the bullet is stuck in the wood *greater than, equal to,* or *less than* the initial kinetic energy of the bullet? ••
17. A block slides with speed v across a low-friction, horizontal surface and collides with and sticks to an identical block at rest. The combination then strikes a spring attached to a fixed wall, bouncing off elastically. (*a*) Choose a system and discuss whether it is isolated. (*b*) Sketch energy bar diagrams for the initial state, the state after the first collision, the state during the collision with the spring, and the state after the final collision. (*c*) Discuss which types of internal energy are involved during this entire sequence of events. (*d*) Determine the final speed of the combined blocks. ••

5.4 Closed systems

18. Light a matchstick by striking at the edge of the match box. If the event takes place in a room that can be considered a closed system, what are the consequences? If the same event occurs inside an isolated system, will there be any change? •
19. Why are most of the massive meteoroids not able to hit the Earth's surface? Explain your answer in terms of energy. •
20. Imagine making two springy devices, each made up of a dozen or so metal blocks loosely connected by springs, and then making the two collide head-on. Do you expect the collision to be elastic, inelastic, or totally inelastic? If the collision is not elastic, where does the kinetic energy go? ••
21. Car A is at rest when an identical car, B, rear-ends it and they move together. What fraction of the kinetic energy of car B is lost? ••
22. In the past, cars were built to be fairly rigid in collisions. Now "crumple zones" (areas designed to crumple or deform during a collision) are deliberately engineered. Why do such zones make a car safer in a collision? ••

5.5 Elastic collisions

23. A ball of inertia m bounces elastically from the floor. It has x component of velocity $+v$ just before it hits the floor. What is the magnitude of its change in momentum in the collision? What is the change in its kinetic energy? Are your two answers consistent with each other? •
24. A 1200-kg car initially at rest undergoes constant acceleration for 8.8 s, reaching a speed of 10 m/s. It then collides with a stationary car that has a perfectly elastic spring bumper. What is the final kinetic energy of the two-car system? •
25. An elastic collision takes place between a 0.080-kg toy car moving at 10×10^2 m/h and a 0.016-kg toy car moving at 20×10^9 mm/yr. What is the kinetic energy of the system in joules? •
26. You want to give a third-grader a common example that illustrates how much energy a joule is. Show that a 0.1-kg ring of keys falling out of your pocket develops about 1 J of kinetic energy before hitting the ground. Can you think of another example? ••
27. You have an inertia of 52 kg and are standing at rest on an iced-over pond in your skates. Suddenly, your 60-kg brother skates in from the right with a speed of 5.0 m/s and collides elastically with you. (*a*) What is the siblings' relative speed after the collision? (*b*) If your brother's final velocity is 0.36 m/s toward the left, what is your final velocity? (*c*) Is your answer to part *b* consistent with your answer to part *a*? (*d*) What is the change in kinetic energy of the system of you and your brother? ••
28. Block A of inertia 2 kg moving at 5 m/s collides elastically with a stationary block, B, of unknown inertia. After the collision, A is moving in the same direction at 3 m/s. What is the inertia of block B? What is its velocity after the collision? ••
29. An object of inertia m_1 moving at v_{1i} collides elastically with an object of inertia m_2 initially at rest. Show that the following solution for the final velocities of the two objects after the collision conserves momentum and kinetic energy:

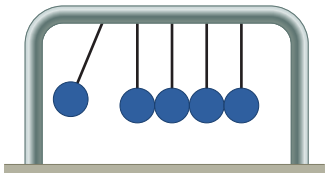
$$v_{1f} = \left(\frac{m_1 - m_2}{m_1 + m_2}\right) v_{1i}$$

$$v_{2f} = \left(\frac{2m_1}{m_1 + m_2}\right) v_{1i}$$ ••

30. Use the solutions in Problem 29 to explain the outcome of each of the following collisions:
    (*a*) A moving truck colliding with a bike initially at rest.
    (*b*) A bike colliding with a stationary truck.
    (*c*) A bike colliding with an identical stationary bike. ••

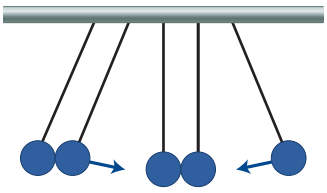
31. The device shown in Figure P5.31 is called a Newton's cradle. Any collision between the balls in the device in Figure P5.31 is elastic. (*a*) If one end ball is pulled back and let go, as shown, would it be a violation of conservation of momentum if two balls at the other end bounced up at half the first ball's speed? (*b*) Would the result in part *a* be a violation of conservation of energy? (*c*) Show that the *only* result consistent with both conservation laws is that for n balls raised and let go, n balls rise on the other side. ●●

Figure P5.31



32. Figure P5.32 shows a pattern for the balls in a Newton's cradle just before collision. The arrows indicate the directions of motion of the balls. All the balls have the same size and inertia, and they collide elastically. Assume that all the collisions happen at the same instant. Sketch the pattern of the balls immediately after the collision. ●●

Figure P5.32



33. A cart of inertia m_1 and velocity $\vec{v}_{1i}$ collides elastically with a cart of inertia m_2 initially moving at velocity $-0.5\vec{v}_{1i}$. What is the speed of each cart after the collision? ●●●
34. Two carts are initially moving to the right on a low-friction track, with cart 1 behind cart 2. Cart 1 has a speed twice that of cart 2 and so moves up and rear-ends cart 2, which has twice the inertia of cart 1. What is the speed of each cart right after the collision if the collision is elastic? ●●●

5.6 Inelastic collisions

35. A Labrador retriever dog with inertia 25 kg and velocity 6 m/s jumps into a trolley at rest and after that both move in the same direction with a velocity 3 m/s. Is the collision elastic, inelastic, or totally inelastic? Calculate the inertia of the trolley. ●
36. A physics student driving a 1200-kg car runs into the rear of a 2000-kg car stopped at a red light. From the fact that the joined vehicles skidded forward 4.0 m, the investigating police officer calculates that the speed after the collision was 6.6 m/s. Based on this information, calculate the student's speed before impact. ●
37. A basketball is thrown horizontally at a heavy door that is free to move. For the two cases in which the ball is (*a*) well inflated and (*b*) flat, sketch a system diagram of the ball and the door, showing momentum vectors and energy bars just before and just after the collision. Make your picture the view from above. ●
38. Draw a velocity-versus-time graph for a collision between a 2-kg object moving at an initial speed of 4 m/s and a 6-kg object initially at rest, if the coefficient of restitution is 0.25. ●●
39. Consider four identical cars. Two of them collide head-on when both are traveling at 34 m/s. The other two collide head-on when both are traveling at 25 m/s. What is the ratio of the amount of kinetic energy converted to internal energy in the two collisions? ●●
40. Take the common case where an object of inertia m_1 collides totally inelastically with a stationary object of inertia m_2. Show that the fraction of kinetic energy converted in the collision is $m_2/(m_1 + m_2)$. Comment on the amount of energy converted for the cases $m_2 \gg m_1$. and $m_2 \ll m_1$, ●●
41. Show that when a moving object collides with an identical stationary object, the ratio of the final and initial kinetic energies is related to the coefficient of restitution e by

$$\frac{K_f}{K_i} = \tfrac{1}{2}(1 + e^2).$$ ●●

42. A 2000-kg truck is sitting at rest (in neutral) when it is rear-ended by a 1000-kg car going 25 m/s. After the collision, the two vehicles stick together. (*a*) What is the final speed of the car-truck combination? (*b*) What is the kinetic energy of the two-vehicle system before the collision? (*c*) What is the kinetic energy of the system after the collision? (*d*) Based on the results of parts *b* and *c*, what can you conclude about which type of collision this is? (*e*) Calculate the coefficient of restitution for this collision. Is this the result you expect for *e*? ●●
43. Which head-on collision between a small car and a large truck causes a larger conversion of kinetic energy: one in which their initial momenta have the same magnitude, or one in which their initial kinetic energies are the same? Assume the same kinetic energy for the two-vehicle system in both cases. ●●●
44. You push a wooden block and friction causes it to stop eventually. If the acceleration during the process is constant, calculate the fraction of its kinetic energy lost as a function of the distance it was pushed through. Check your answer at the middle and end of its path. ●●●

5.7 Conservation of energy

45. An experienced bartender knows just the right initial speed of a glass of beer to get it to come to a stop in front any customer sitting along the bar. Say the initial speed needed to move a glass all the way to the end of the bar is v_{end}. In terms of v_{end}, with what speed should an identical glass of beer be released if it is to stop at a customer sitting three-quarters of the way down the bar? (The kinetic energy converted to thermal energy because of friction is proportional to the distance the glass skids.) •
46. A car goes into a skid and gradually comes to a stop, accelerating at a constant rate. At the midpoint of the skid, how much of its kinetic energy has it lost? •
47. A 2-kg block moving at 4 m/s hits a slower 1-kg block moving at 2 m/s. If the 2-kg block continues in the same direction at 3 m/s, (*a*) what is the speed of the slower block? (*b*) Is the collision elastic? (*c*) By what factor does the internal energy of the system change? ••
48. A 1200-kg car is backing out of a parking space at 5.0 m/s. The unobservant driver of a 1800-kg pickup truck is coasting through the parking lot at a speed of 3.0 m/s and runs straight into the rear bumper of the car. (*a*) What is the change in internal energy of the two-vehicle system if the velocity of the pickup is 1.5 m/s backward after they collide? (*b*) What is the coefficient of restitution for this collision? ••
49. Explain how a police officer at the scene of a car crash can judge whether or not you were speeding when you rear-ended a stopped car, based solely on the length of the skid marks after the collision. Is the speed the officer estimates for your speed just before the collision faster than, equal to, or slower than your actual speed? ••
50. A wagon is coasting along a level sidewalk at 5.00 m/s. Its wheels have very good bearings. You are standing on a low wall, and you drop vertically into the wagon as it passes by. The wagon has an inertia of 100 kg, and your inertia is 50.0 kg. (*a*) Use conservation of momentum to determine the speed of the wagon after you are in it. (*b*) Use conservation of energy to determine that speed. (*c*) After comparing your answers in parts *a* and *b*, explain which method is correct. ••
51. The extinction of the dinosaurs is, by one theory, attributed to a collision between Earth and an asteroid about 10 km in diameter moving at 25 km/s relative to Earth. Assume that the asteroid had about the same density as Earth. Estimate the energy released by such an impact in megatons of TNT, where 1 megaton $= 4.2 \times 10^{15}$ J. •••

5.8 Explosive separations

52. Can you tell from the coefficient of restitution whether a collision has added kinetic energy to a system, taken some away, or left the system's kinetic energy unchanged? •
53. If the video of an explosive separation (like the firing of a gun) is run backward, the event looks like a totally inelastic collision, with two or more originally separate pieces all sticking together. How are the coefficient of restitution for an explosive separation and that for a totally inelastic collision related to each other? •
54. At the peak of its vertical flight, a fireworks shell explodes into two pieces. Piece 1 has three times the speed of piece 2. What is the ratio of the inertias of the pieces? What is the ratio of their kinetic energies? ••
55. A 52-kg ice skater (this value includes her body, her clothing, and several 1.0-kg snowballs she is carrying) is at rest on the ice. She throws a snowball to the right at 10 m/s. (*a*) What is her speed after the throw? Is her velocity to the left or to the right? (*b*) Calculate the coefficient of restitution for this event. She next throws a second snowball but this time at a speed of 20 m/s to the left. (*c*) What is her speed after this throw? Is her velocity to the left or to the right? (*d*) Calculate the coefficient of restitution for this event. (*e*) Calculate the change in kinetic energy in this second event. Where does the added kinetic energy come from? (*f*) If one food calorie equals 4184 J, how many food calories does the skater burn when she throws the second snowball? ••
56. A mysterious crate has shown up at your place of work, Firecracker Company, and you are told to measure its inertia. It is too heavy to lift, but it rolls smoothly on casters. Getting an inspiration, you lightly tape a 0.60-kg iron block to the side of the crate, slide a firecracker between the crate and the block, and light the fuse. When the firecracker explodes, the block goes one way and the crate rolls the other way. You measure the crate's speed to be 0.055 m/s by timing how long it takes to cross floor tiles. You look up the specifications of the firecracker and find that it releases 9.0 J of energy. That's all you need, and you quickly calculate the inertia of the crate. What is that inertia? ••
57. A mysterious explosion in an unmanned 6000-kg spaceship coasting at 5000 m/s breaks the ship into two unequal parts, one twice as massive as the other. The smaller piece recoils in the direction opposite to the original path at 3000 m/s. (*a*) What is the velocity of the larger piece after the explosion? (*b*) Discuss the energy changes resulting from the explosion. ••
58. A system consists of a 4.0-kg cart and a 1.0-kg cart attached to each other by a compressed spring. Initially, the system is at rest on a low-friction track. When the spring is released, an explosive separation occurs at the expense of the internal energy of the compressed spring. If the change in the spring's internal energy during the separation is 1.0 kJ, what is the speed of each cart right after the separation? ••
59. A space shuttle of inertia m is attached to a booster rocket that has an inertia nine times greater. This system is moving at a speed of 800 m/s in outer space, as seen

by observers in a nearby space station. Then explosive bolts are detonated, separating the shuttle from the rocket and thrusting the shuttle forward at a speed of 100 m/s relative to the rocket. What are the velocities of the rocket and the shuttle right after the explosion? ••

60. A uranium-238 atom can break up into a thorium-234 atom and a particle called an *alpha particle*, α-4. The numbers indicate the inertias of the atoms and the alpha particle in *atomic mass units* ($1\text{ amu} = 1.66 \times 10^{-27}\text{ kg}$). When a uranium atom initially at rest breaks up, the thorium atom is observed to recoil with an x component of velocity of -2.5×10^5 m/s. How much of the uranium atom's internal energy is released in the breakup? ••

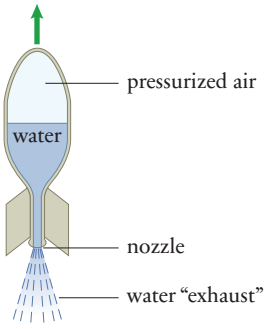
61. A stationary 5.0-kg ball explodes into two pieces due to a firecracker inside. One piece with inertia 2.0 kg travels to the left at 10 m/s. (*a*) What is the speed of the second fragment? (*b*) How much energy was released in the explosion? ••

62. An assembled system consists of cart A of inertia m_A, cart B of inertia m_B, and a spring of negligible inertia, clamped together so that the fully compressed spring is aligned between the front end of cart B and the back end of cart A. The internal energy of the system, stored in the compressed spring, is E_{spring}. The system is put on a low-friction track and given a small shove so that it moves to the right at speed v_i, with cart A in front. Once the system is moving, the clamp is released (by remote control, so that the motion of the system as a whole is not affected in any way). As the spring expands, the carts move apart. What are their final speeds? •••

63. Consider the same initial system of spring and carts A and B as in Problem 62. After the system is set in motion with common speed v_i, the clamp is again released. What is the speed of cart A at the instant when the spring is still partially compressed and stores one-fourth of its initial internal energy E_{spring}? •••

64. A toy water-rocket consists of an elongated plastic tank with a nozzle at one end (Figure P5.64). A hand pump is used to get pressurized air into the tank. If there is nothing but air in the tank, the performance is lackluster. If some water is put in the tank before it is pumped up with air, however, the flight is dramatically better. Why? •••

Figure P5.64



65. A fictional subatomic particle, the solarino, is initially at rest. It then explodes into three particles: two lightons and a heavyon. The inertia of a heavyon is twice the inertia of a lighton. In the explosive separation, the three particles can be kicked out in either direction along a straight line, but the internal energy of the solarino is always entirely converted to the kinetic energy of the three particles. (*a*) Draw the paths of the three particles for the case in which the heavyon has the greatest possible speed. (*b*) Draw the paths of the particles for the case in which one of the lightons has the greatest possible speed. •••

Additional Problems

66. Which has more kinetic energy: A 163-g cricket ball travelling at 150 km/h or a 58-g tennis ball travelling at 230 km/h? •

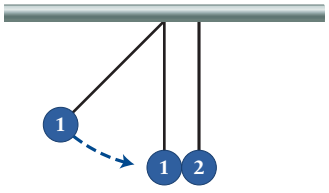
67. In javelin throw, an athlete usually runs a certain distance before the throw. Why does the athlete not throw their javelin from a stand-still position? •

68. You and a friend are playing catch with a medicine ball. The game has gotten a little unfriendly, and you'd like to knock your friend down with the next throw. To achieve this, should you suggest that your friend catch the ball with his hands or bounce it elastically back to you with his fists? ••

69. Your uncle works at the railroad freight yard, and he has asked you to watch the yard hands couple 20,000-kg boxcars together. As he leaves to get a sandwich, he tells you, "Whatever you do, don't let those cars bang together too hard. The couplings will take only 10,000 J on a hit before they shatter." You ask how hard is too hard, but your uncle has already ducked out. How fast can one moving car bump into a stationary (but free-to-move) car without the coupling shattering? ••• CR

70. Two solid spheres hung by thin threads from a horizontal support (Figure P5.70) are initially in contact with each other. Sphere 1 has inertia $m_1 = 0.050$ kg, and sphere 2 has inertia $m_2 = 0.10$ kg. When pulled to the left and released, sphere 1 collides elastically with sphere 2. At the instant just before the collision takes place, sphere 1 has kinetic energy $K_1 = 0.098$ J. (*a*) What is the velocity of sphere 1 right before the collision? (*b*) What is the kinetic energy of the system before the collision? (*c*) What is the velocity of each sphere after the collision? (*d*) From part *c*, calculate the kinetic energy after the collision. Does the value you get equal the result of part *b*? Explain why or why not. (*e*) Calculate the coefficient of restitution of the collision. Is this the result you expect? ••

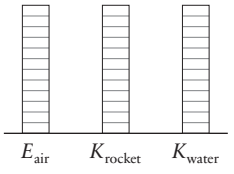
Figure P5.70



71. Repeat Problem 70 for two spheres made of modeling clay. All the data remain the same, but this time the two spheres stick together upon collision. ••

72. In the toy water-rocket of Problem 64, the pressurized air can store internal energy. Assume you have a rocket that expels water at a speed of 15.0 m/s and the opening through which the water leaves is 10.0 mm in diameter. (*a*) Using Figure P5.72 as a template, sketch energy bar graphs for three instants: when the rocket is fully pressurized but has not been launched, after the launch when half the water has been expelled, and after all the water has been expelled. (*b*) What is the approximate inertia of the water expelled in the first second of flight? (*c*) Does the rocket accelerate at a constant rate? ••

Figure P5.72



73. Suppose that, as it slides across the floor of a basketball arena, a hockey puck loses kinetic energy because of friction. The amount of energy lost is directly proportional to the distance the puck slides. Sketch a graph of kinetic energy versus time during this slide and then verify with a calculation. ●●●

74. While bowling one day, you begin to wonder about the inertia of a bowling pin. You know that balls and pins generally make elastic collisions and that your 6.5-kg bowling ball continues to move forward after a collision, even when you hit several pins to make a strike. As it happens, you have just managed to knock down nine of the ten pins, leaving only one pin upright at the end of the lane. You ask your friend to make a video of your next shot so that you can later carefully observe the initial and final velocities of your ball. Sure enough, you hit the remaining pin head-on (make the spare), and after a little video analysis you decide that your ball lost about 40% of its initial speed when it hit the pin. What is the inertia of that pin? ●●● CR

75. For a system of two identical cars, A and B, of the same inertia m but moving at two different velocities, v_A and v_B, show that the kinetic energy of the two-car system can be expressed as the sum of two terms: the kinetic energy of a double car moving with one-half the sum of their velocities, plus the kinetic energy of a double car moving with one-half the difference of their velocities:

$$K_{sum} = \tfrac{1}{2}(2m)\left[\frac{v_A + v_B}{2}\right]^2$$

$$K_{dif} = \tfrac{1}{2}(2m)\left[\frac{v_A - v_B}{2}\right]^2.$$

Interpret the meaning of each of these terms. ●●●

76. A 1000-kg car traveling with an x component of velocity of +20 m/s collides head-on with a 1500-kg light truck traveling with an x component of velocity of −10 m/s. (*a*) If 10% of the system's kinetic energy is converted to internal energy during the collision, what are the final speeds of the car and truck? (*b*) If the car had instead rear-ended the same truck moving with an x component of velocity of +10 m/s, how would your answer to part *a* change? ●●●

77. Your little brother is outside squirting a basketball with the stream of water from a garden hose. As the ball rolls across the yard, you begin to wonder what you would need to know in order to determine the rate at which momentum is transferred from the water to the ball. It should not be hard to measure both the radius r of the hose and the flow rate Q of water out of the hose in cubic meters per second. The density of water is certainly known, but you aren't sure whether the collision is elastic or inelastic. ●●● CR

78. Two hockey players push off from a clinch, recoiling from rest. Jean-Claude has an inertia that is 50% greater than Pierre's inertia. After they push off and move across the ice in opposite directions, they lose kinetic energy at the same rate until they come to a stop. (*a*) Who travels farther? (*b*) How much farther? ●●●



6

Principle of Relativity

CONCEPTS

6.1 Relativity of motion
6.2 Inertial reference frames
6.3 Principle of relativity
6.4 Zero-momentum reference frame

QUANTITATIVE TOOLS

6.5 Galilean relativity
6.6 Center of mass
6.7 Convertible kinetic energy
6.8 Conservation laws and relativity

Did you ever, sitting in a car at a red light and looking at the car next to you, slam on the brakes because you thought you were starting to roll but it turned out your car never moved? If you did, you experienced the relativity of motion, first described quantitatively by Galileo. You may have heard the term *relativity* in the context of Albert Einstein's famous theories of general relativity and special relativity. We'll touch upon one basic aspect of general relativity in Chapter 13 and discuss special relativity in Chapter 14. Underlying all of relativity is something we have so far ignored in our discussion of motion: the fact that the velocity measured for any object depends on the **observer** (the person doing the measuring). For example, to a person sitting in a moving train, a suitcase on the overhead rack is at rest, but to a person standing on a station platform watching the train speed by, the suitcase is not at rest. According to the person in the train, the suitcase has zero momentum $(m\vec{v})$ and zero kinetic energy $\left(\frac{1}{2}mv^2\right)$. According to the person on the platform, the suitcase has nonzero momentum and nonzero kinetic energy. In this chapter we investigate whether or not the laws of conservation of momentum and conservation of energy depend on the velocity of the observer. In other words, if these laws are valid for one observer, are they also valid for an observer who is moving relative to the first observer?

6.1 Relativity of motion

Remember from Chapter 2 that whenever we talk about motion, we must specify a reference axis along which the motion occurs and an origin. Together, the axis and origin are called a **reference frame.** In our study of motion thus far, we have considered the motion of carts along a low-friction track, collecting our data using a ruler affixed to the track; that is to say, we collected our data while standing in a reference frame at rest relative to the surface of Earth (this is called an **Earth reference frame**).

What happens to the data if we are moving as we collect them—in other words, if we are in a reference frame that is moving relative to Earth? Imagine that we are moving alongside the track and want to describe a collision between two carts. Do the conservation laws hold in a reference frame that is moving relative to Earth? Or, to put the question another way, does the behavior of the carts change if we are in motion as we study that behavior?

Consider, for example, the situation depicted in **Figure 6.1*a***: two observers watching the motions of two cars. Car 1 is moving at constant velocity $\vec{v}_1$, and car 2 is at rest relative to Earth. Observer E is at rest relative to Earth, while observer M is inside car 1. The velocity $\vec{v}_M$ of observer M is the same as the velocity of car 1: $\vec{v}_M = \vec{v}_1$.

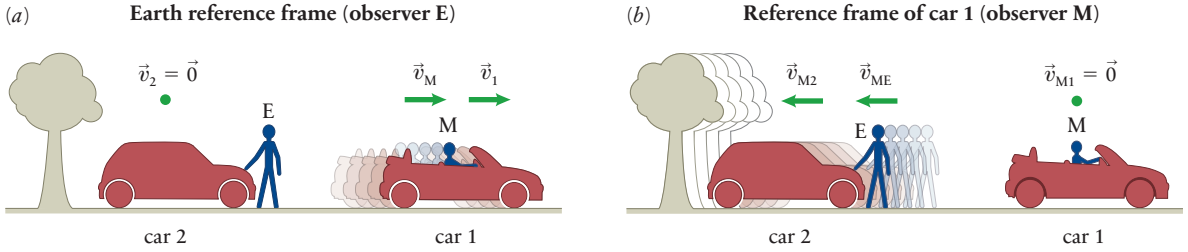
Relative to observer M, the position of car 1 is not changing; that is, car 1 is at rest relative to observer M. The distance to car 2, however, is steadily increasing, and so relative to observer M, car 2, observer E, and all the surroundings move to the left (Figure 6.1*b*). To distinguish quantities measured by different observers, we add a capital letter subscript to denote the observer. For instance, $\vec{v}_{M2}$ represents observer M's measurement of the velocity of car 2, also called the velocity of car 2 *relative to* observer M. When we are comparing quantities in various reference frames, quantities measured relative to Earth are given the subscript E. For example, relative to Earth the velocity of car 1 is represented by $\vec{v}_{E1}$ (When the Earth reference frame is the only reference frame, we drop the subscript E, recovering the notation we have used before: $\vec{v}_{E1} = \vec{v}_1$.)

To determine a relationship between the quantities measured by two observers moving relative to each other, let's start with a simple example. **Figure 6.2** on the next page shows five frames of a video of two carts on a low-friction track. In Figure 6.2*a* the positions of the two carts are measured with ruler A, which is affixed to the track (that is, in the Earth reference frame). Relative to ruler A, cart 1 is at rest and its left end remains fixed at 22.5 mm. Cart 2 is moving to the right, and the position of its right end on ruler A changes from 0 to 14.4 mm.

In Figure 6.2*b* the positions of the two carts are measured with ruler B, which moves along with cart 2: The zero mark on ruler B remains aligned with the front of cart 2. The position of cart 2 therefore remains fixed at

CONCEPTS

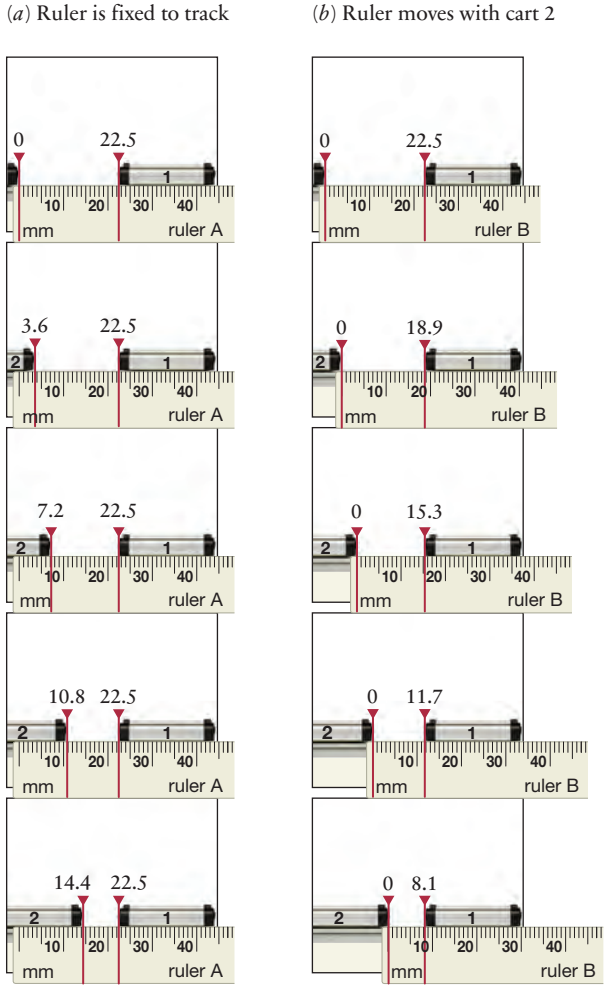
Figure 6.1 The motion you perceive depends on your reference frame.



Relative to observer E, car 2 is at rest and car 1 moves to the right.

But relative to observer M, car 1 is at rest while car 2, observer E, and Earth move to the left.

Figure 6.2 Two identical carts on a low-friction track. The positions of the carts are measured with (*a*) a ruler affixed to the track and (*b*) a ruler moving along with cart 2. In the Earth reference frame, cart 1 is at rest and cart 2 is moving to the right.



a ruler reading of 0, and so, relative to ruler B, cart 2 is at rest. On this ruler, the position of cart 1 does *not* remain fixed, however—the reading for this cart's position decreases from 22.5 mm to 8.1 mm over the five frames. That is to say, cart 1 moves 22.5 mm − 8.1 mm = 14.4 mm *to the left* along ruler B.

Figure 6.3*a* shows a position-versus-frame graph for the two carts drawn from the readings on ruler A in Figure 6.2*a*. What is shown in the graph corresponds to what an observer at rest alongside the track sees: cart 1 at rest and the x component of the velocity of cart 2 a constant +3.6 mm/frame. These are the observations taken in the Earth reference frame.

Figure 6.3*b* is the position-versus-frame graph for the two carts drawn from the readings on ruler B in Figure 6.2*b*. What is shown in the graph corresponds to what an observer M sitting on top of cart 2 sees: cart 2 at rest and the x component of the velocity of cart 1 a constant −3.6 mm/frame (the negative sign indicates that the cart moves to the left in the figure).

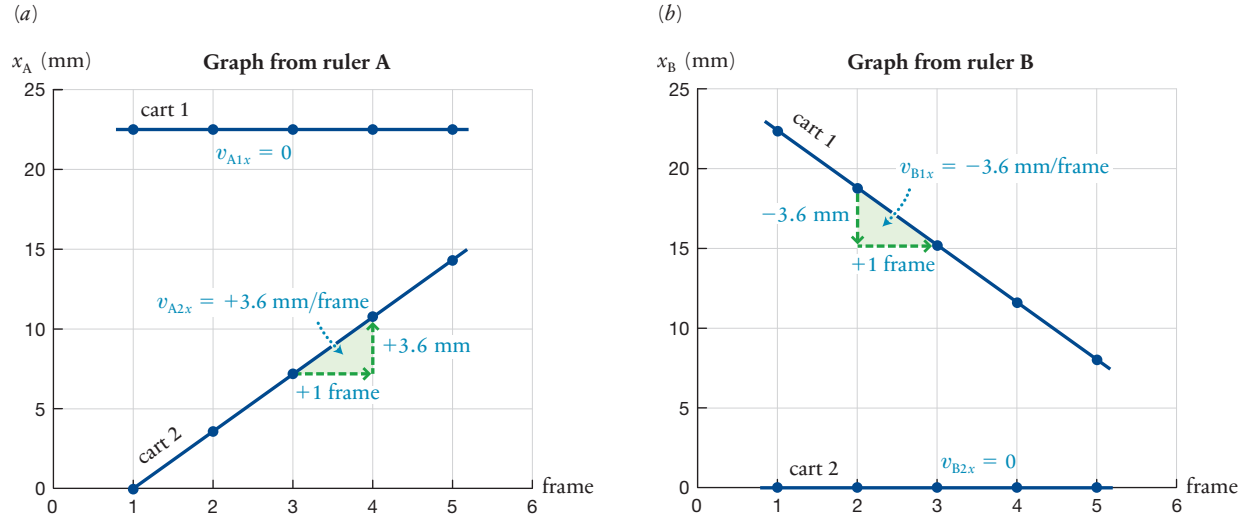
EXAMPLE 6.1 Moving at my own speed

Suppose that in the situation shown in Figure 6.2, a third ruler is affixed to some device (not shown) that moves to the right along the track at a speed of 2.0 mm/frame. If in the Earth reference frame cart 2 again moves at +3.6 mm/frame and cart 1 is again at rest, what are the cart velocities according to an observer moving along with ruler C?

1 GETTING STARTED If I arbitrarily let ruler C start in the position shown in frame 1 of Figure 6.2*b*—the zero mark on ruler C aligned with the front of cart 2—this ruler is shifted 2.0 mm to the right in the second frame and another 2.0 mm to the right in each successive frame. I

CONCEPTS

Figure 6.3 Position-versus-frame graphs for the carts of Figure 6.2 as determined from readings on (*a*) ruler A and (*b*) ruler B.



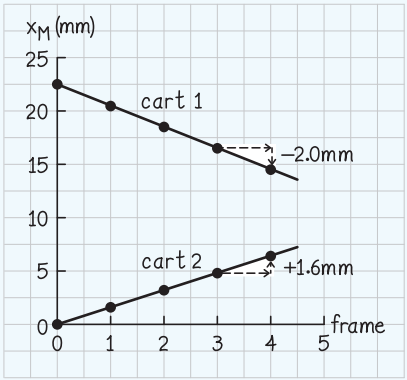
have to use the readings from this ruler to determine the cart velocities.

2 DEVISE PLAN I begin by making a table showing the position of each cart in each frame according to an observer moving along with ruler C. Because the ruler moves 2.0 mm to the right in each frame, I can obtain the readings on this ruler by subtracting 2.0 mm from each cart position in frame 2 of Figure 6.2*a*, subtracting 4.0 mm from each cart position in frame 3, subtracting 6.0 mm from each cart position in frame 4, and so on. I can then use these positions to make a position-versus-frame graph like the ones in Figure 6.3 and determine the velocities of the carts from the slopes of the curves.

3 EXECUTE PLAN My table of the cart positions measured with ruler C is shown in **Figure 6.4**. By plotting these data, I see that the x component of the velocity of cart 1 is -2.0 mm/frame and that of cart 2 is $+1.6$ mm/frame. ✔

Figure 6.4

x_{M2} (mm)	x_{M1} (mm)
0	22.5
1.6	20.5
3.2	18.5
4.8	16.5
6.4	14.5



4 EVALUATE RESULT The observer moving with ruler C sees cart 1 moving to the left, meaning that the x component of its velocity should be negative, as I found. This observer and cart 2 both move in the positive x direction in the Earth reference frame, but the speed of cart 2 is higher than that of ruler C, and so cart 2 moves to the right relative to the observer. Therefore the x component of its velocity should be positive, in agreement with what I found. Because the velocity of ruler C is different from the velocity of either cart, neither cart is at rest relative to ruler C.

This example and the text preceding it show that motion is a relative concept. In the Earth reference frame, cart 1 is at rest and cart 2 is moving. Relative to the moving ruler in Figure 6.2*b*, cart 1 is moving and cart 2 is at rest. And relative to the ruler of Example 6.1, both carts are moving. In all three reference frames, however, we discover that the carts move at constant velocity. We could just as well have used a ruler moving at $+1.0$ mm/frame, $+4.0$ mm/frame, -0.1 mm/frame, or any other constant velocity. In each case, we would reach the same conclusion: The carts move at constant velocity. So, if an object moves at constant velocity in the Earth reference frame, its motion observed from *any reference frame moving at constant velocity relative to Earth* is also at constant velocity.

Comparing Figures 6.3*a* and 6.3*b*, we see that the x component of each cart's velocity according to ruler A (Figure 6.3*a*) is 3.6 mm/frame higher than the velocity according to ruler B (Figure 6.3*b*). So, an object's velocity $\vec{v}_{Ao}$ determined from the readings on ruler A is equal to the sum of the velocity $\vec{v}_{AB}$ of ruler B relative to ruler A and the velocity $\vec{v}_{Bo}$ of the object relative to ruler B: $\vec{v}_{Ao} = \vec{v}_{AB} + \vec{v}_{Bo}$. For cart 1, we have $0 = +3.6$ mm/frame $+ (-3.6$ mm/frame); for cart 2, $+3.6$ mm/frame $= +3.6$ mm/frame $+ 0$. You can readily verify that the same is true for the velocities determined with ruler C in Example 6.1.



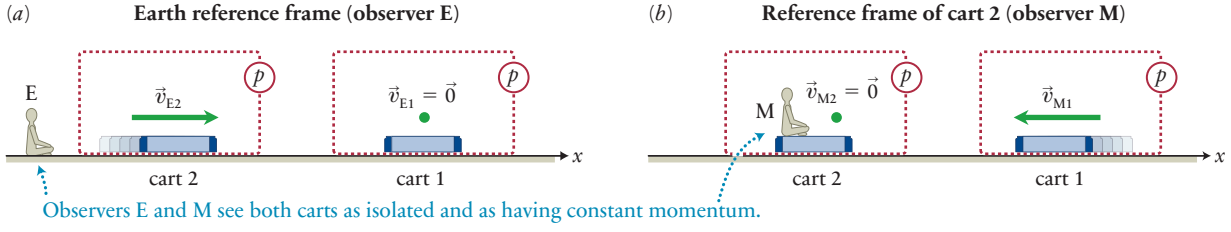
6.1 What is the velocity of each cart in Figure 6.2 measured by an observer moving at -3.0 mm/frame in the Earth reference frame?

6.2 Inertial reference frames

How does the choice of reference frame affect our accounting methods for momentum and energy? To answer this question, let's consider the two carts of Figure 6.2 from the point of view of two observers. To an observer E, who is at rest relative to Earth, cart 1 is at rest and cart 2 is moving at constant velocity (**Figure 6.5a**). To an observer M, who is moving along with cart 2, cart 2 is at rest and cart 1 is moving at constant velocity (Figure 6.5*b*). Even though the observers obtain different values for the carts' velocities, these

CONCEPTS

Figure 6.5 The carts of Figure 6.2 seen by (*a*) an observer in the Earth reference frame and (*b*) an observer moving along with cart 2.



values are constant, and therefore both observers conclude that the momentum of each cart is constant. Both observers also agree that the two carts are isolated, which means that conservation of momentum yields $\Delta\vec{p} = \vec{0}$ (Eq. 4.17) for each cart, in agreement with their observations.

6.2 From the point of view of each observer in Figure 6.5, (*a*) is the energy of each cart constant? (*b*) Is the isolated system containing cart 1 closed? (*c*) Is the isolated system containing cart 2 closed?

Although we checked for only the simple case of carts moving at constant velocity on a low-friction track, Figure 6.5 and Checkpoint 6.2 suggest that the accounting procedures and principles we introduced in the preceding two chapters can be applied in any reference frame moving at constant velocity relative to Earth. Indeed, when we are moving at constant velocity relative to Earth, things around us behave the same way they behave when we are at rest on the ground. For example, imagine being in a plane flying at a constant speed of 260 m/s in a straight line. The surface of the coffee in the cup in front of you looks no different from the way it looks at home. If you drop your keys, they fall straight down just as they would if the plane were at rest on the ground. And when the flight attendant pours coffee into your cup while you all move at 260 m/s, you don't expect him to spill it on you any more than you do when a waiter pours your coffee in a restaurant.

In fact, if the engines are very quiet, the ride is smooth, the cabin is closed, and the window shades are down, it is impossible for you to determine whether or not the plane is moving. Here is the simplest possible experiment: Put a marble on the floor of the plane. Provided the floor is level relative to the ground, the marble remains at rest where you placed it. This is true whether the plane is at rest or cruising at a constant 260 m/s. When the plane is sitting at the gate, inertia keeps the marble at rest relative to the floor (which is at rest relative to the ground), and so $\vec{a} = \vec{0}$. In the air, inertia keeps the marble moving along with the plane at a constant velocity relative to the ground (or, put differently, inertia keeps the marble at rest relative to the floor), and so again $\vec{a} = \vec{0}$.

The Earth reference frame or any reference frame moving at constant velocity relative to Earth is called an **inertial reference frame.** We can tell whether or not a reference frame is inertial by testing whether or not the **law of inertia** holds:

> **In an inertial reference frame, any isolated object that is at rest remains at rest, and any isolated object in motion keeps moving at a constant velocity.**

The law of inertia does not hold in a reference frame that is accelerating relative to Earth. For example, imagine sitting in the passenger seat of a car moving at constant velocity with a marble on the floor near your feet. As long as the car is on a horizontal surface and keeps going straight at a constant speed, the marble remains at rest. When the driver suddenly accelerates forward, however, everything in the car lurches backward, with the marble probably ending up somewhere in the back of the car. According to an observer at rest in the Earth reference frame (**Figure 6.6*a***), the marble resists being accelerated forward because of its inertia, and unless you have glued it to the floor or restrained it in some other way, it fails to keep up with the accelerating car. Seen from inside the car, the isolated marble suddenly accelerates backward, and so the law of inertia does not hold in the reference frame of the accelerating car (Figure 6.6*b*). Reference frames in which the

CONCEPTS

Figure 6.6 Motion of a marble in an accelerating car viewed from (*a*) the Earth reference frame and (*b*) a reference frame affixed to the car. $\vec{a}_{Ec}$ is the Earth observer's measurement of the car's acceleration, $\vec{a}_{ME}$ is the moving observer's measurement of Earth's acceleration, and $\vec{a}_{Mc}$ is the moving observer's measurement of the car's acceleration.

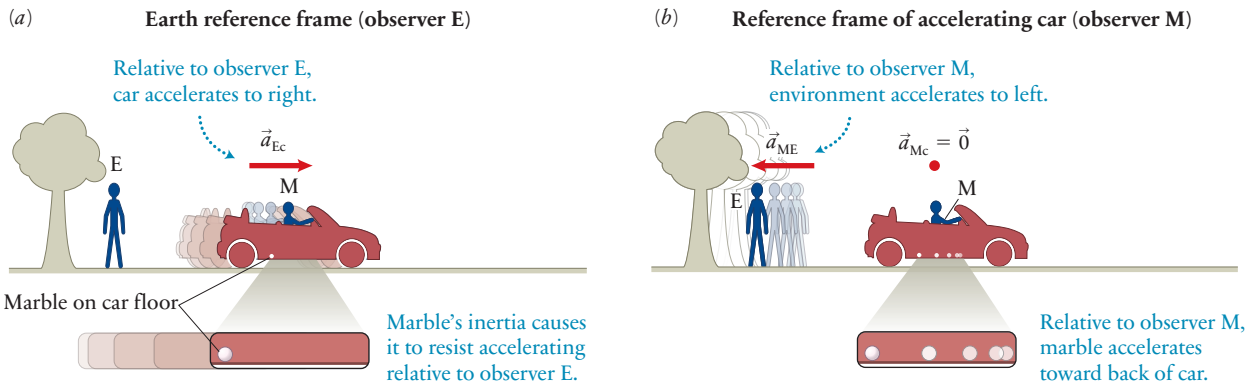
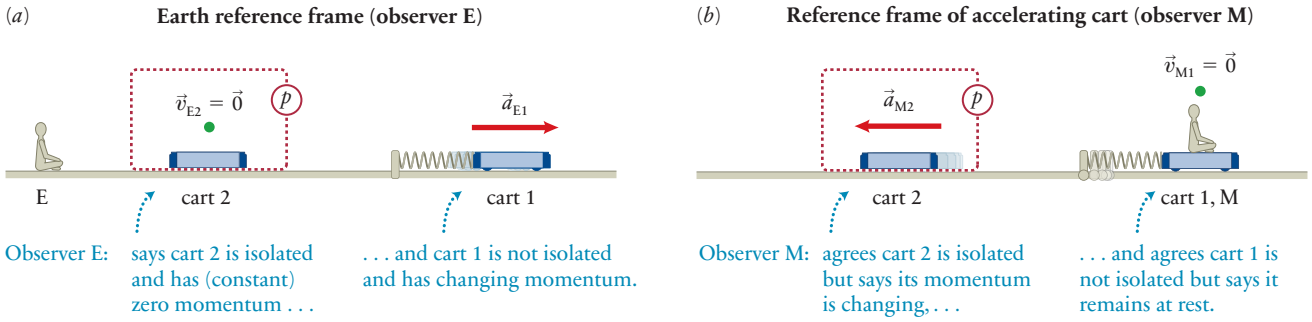


Figure 6.7 A stationary cart and an accelerating cart viewed from (*a*) the Earth reference frame and (*b*) a reference frame affixed to the accelerating cart.



law of inertia does not hold are called *noninertial reference frames*.

We have singled out the Earth reference frame as the basic inertial reference frame. There is nothing fundamentally special about the Earth reference frame, however, other than that we perform most experiments in this frame. Strictly speaking, Earth is not an inertial reference frame because it revolves around a north-south axis and orbits the Sun in a nearly circular orbit. Because motion on a curved path means that the direction of the velocity is changing and so the velocity is not constant, Earth is accelerating and therefore is a noninertial reference frame. For most cases, the acceleration is too small to be noticeable, and so we may, for most practical purposes, consider the Earth reference frame to be inertial.

EXERCISE 6.2 Inertial or not?

Which of these reference frames are inertial: one affixed to (*a*) a merry-go-round, (*b*) the space shuttle orbiting Earth, (*c*) an airplane taking off, (*d*) a train moving at constant speed along a straight track?

SOLUTION (*b*) and (*d*). If I imagine placing a marble at rest in each reference frame and then observing the marble's motion, I can tell whether the law of inertia holds in that reference frame. From experience, I know that a marble remains at rest on the floor of a train moving at constant speed along a straight track, and so the train's reference frame is inertial. I also know that a marble moves away from where I place it on a merry-go-round and in an airplane taking off, so those reference frames are noninertial. Having never been on the space shuttle, I have no direct experience with such a reference frame. However, I have seen astronauts and objects floating in the space shuttle, and this motion suggests that objects stay at rest or keep moving as the law of inertia states, suggesting that the reference frame of the shuttle is inertial.* ✔

*Just like the Earth reference frame, the shuttle reference frame is only approximately inertial.

Our accounting procedures for momentum and energy cannot be used in noninertial reference frames. Consider, for example, the two carts shown in **Figure 6.7**. Cart 1 is being accelerated by a spring, and cart 2 is at rest in the Earth reference frame. Cart 2 constitutes an isolated system, but cart 1 is not isolated because it interacts with the spring. To observer E in the Earth reference frame (Figure 6.7*a*), the behavior of both carts is in agreement with the momentum law: The momentum of the nonisolated cart 1 changes, while the momentum of the isolated cart 2 is constant. For observer M, however, who is accelerating along with cart 1 (Figure 6.7*b*), things don't quite add up. From this observer's perspective, cart 1 remains at rest even though it interacts with the spring, and the momentum of cart 2 changes even though that cart is isolated. Equations 4.17 ($\Delta\vec{p} = \vec{0}$) and 5.23 ($\Delta E = 0$), which embody the laws of conservation of momentum and energy, do not hold in the noninertial reference frame of observer M in Figure 6.7*b*.

Are we going to run into problems because the laws of the universe are different in noninertial reference frames? No, because nothing prescribes the reference frame; we get to choose it. So for now we just avoid using noninertial reference frames. For the accelerating car, for instance, we would choose not a reference frame affixed to the car but the Earth reference frame.

6.3 From the point of view of each observer in Figure 6.7, (*a*) is the energy of each cart constant? (*b*) Is the isolated system containing cart 1 closed? (*c*) Is the isolated system containing cart 2 closed? (*d*) Do the observations made by each observer agree with the conservation of energy law?

6.3 Principle of relativity

In the preceding section, we saw that the conservation laws apply for single objects in inertial reference frames. Let us now test the conservation laws for interacting objects.

CONCEPTS

Figure 6.8 Velocity-versus-time graphs for two carts colliding on a low-friction track as seen (*a*) from the Earth reference frame and (*b*) from a reference frame moving along the track at $v_{EMx} = -0.20$ m/s relative to Earth. The inertias are 0.36 kg for cart 1 and 0.12 kg for cart 2.

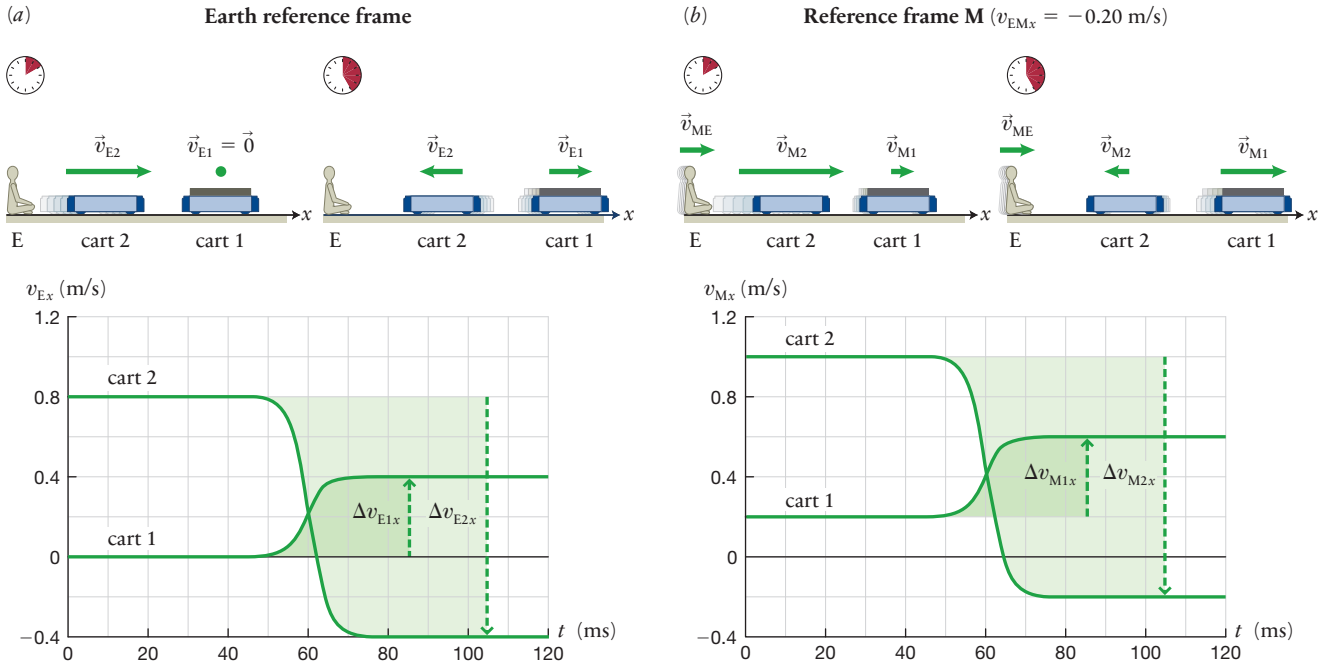


Figure 6.8 shows velocity-versus-time graphs for two-cart collisions. The values in Figure 6.8*a* were measured by an observer in the Earth reference frame. As you saw in Checkpoint 6.1, we can obtain the velocities of the carts in reference frame M of Figure 6.8*b* by subtracting the velocities in reference frame M from the velocities in the Earth reference frame. Because reference frame M in Figure 6.8*b* moves at a constant −0.20 m/s relative to Earth, we must subtract −0.20 m/s (that is, add 0.20 m/s) from the x component of each velocity in the Earth reference frame to obtain the corresponding velocity in reference frame M.

The only difference between Figures 6.8*a* and 6.8*b* is that the curves in Figure 6.8*b* are shifted upward by 0.20 m/s, the relative speed of the two reference frames. However, the *changes* in the x components of the cart's velocities, Δv_{1x} and Δv_{2x}, are not affected by the velocity of reference frame M. Because the inertia of any object is not affected by the motion of an observer studying the object, we observe that the changes in the x component of the carts' momenta, $\Delta p_{1x} = m_1 \Delta v_{1x}$ and $\Delta p_{2x} = m_2 \Delta v_{2x}$, are the same in both reference frames. So, because conservation of momentum requires that $\Delta \vec{p} = \vec{0}$ in the Earth reference frame, we also obtain $\Delta \vec{p} = \vec{0}$ in reference frame M.

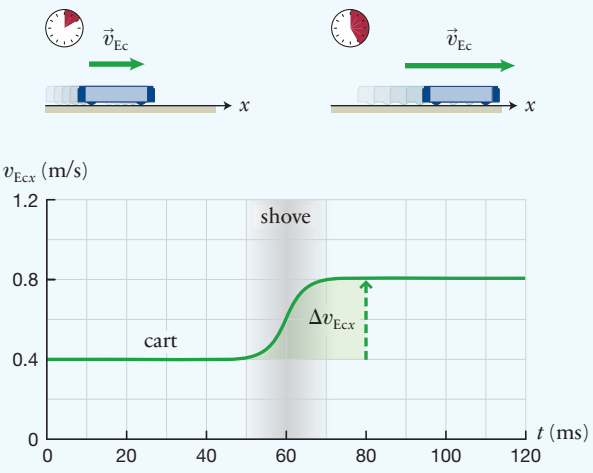
If reference frame M in Figure 6.8*b* had moved at any other constant velocity relative to Earth, the vertical shift of the curves would be different but the shape of the curves would be the same. Again we would obtain the same momentum changes for the two carts, and again we would conclude that the collision does not change the momentum of the system. So, if the change in the system momentum is zero in the Earth reference frame, it must be zero in all reference frames moving at a constant velocity relative to Earth. Experiments show that this statement can be extended to any inertial reference frame, not just the Earth reference frame.

CONCEPTS

EXAMPLE 6.3 Momentum shove

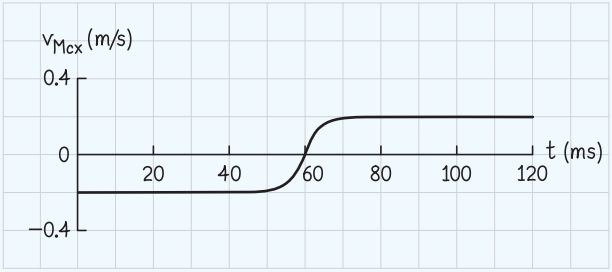
A 0.12-kg cart moving on a straight, low-friction track gets a shove so that its speed changes. Its initial and final speeds measured in the Earth reference frame are 0.40 m/s and 0.80 m/s (**Figure 6.9**). Determine the change in the cart's momentum as seen from (*a*) the Earth reference frame and (*b*) a reference frame M moving in the same direction as the cart at a speed of 0.60 m/s relative to the track.

Figure 6.9 Example 6.3.



1 GETTING STARTED To determine the momentum change in either case, I need to know the cart's initial and final velocities in the two reference frames. In the Earth reference frame, $v_{\mathrm{Ec}x,\mathrm{i}} = +0.40$ m/s and $v_{\mathrm{Ec}x,\mathrm{f}} = +0.80$ m/s. In reference frame M, which moves at $v_{\mathrm{ME}x} = +0.60$ m/s relative to Earth, these values are 0.60 m/s lower: $v_{\mathrm{Mc}x,\mathrm{i}} = -0.20$ m/s and $v_{\mathrm{Mc}x,\mathrm{f}} = +0.20$ m/s. I then make a velocity-versus-time graph for the cart as viewed from reference frame M (**Figure 6.10**).

Figure 6.10



2 DEVISE PLAN Now that I have the x components of the initial and final velocities in both reference frames, I can calculate the x components of the initial and final momenta and determine the change in momentum.

3 EXECUTE PLAN (*a*) In the Earth reference frame, the x components of the initial and final momenta are $p_{\mathrm{Ec}x,\mathrm{i}} = mv_{\mathrm{Ec}x,\mathrm{i}} = (0.12\ \mathrm{kg})(+0.40\ \mathrm{m/s}) = +0.048\ \mathrm{kg \cdot m/s}$ and $p_{\mathrm{Ec}x,\mathrm{f}} = mv_{\mathrm{Ec}x,\mathrm{f}} = (0.12\ \mathrm{kg})(+0.80\ \mathrm{m/s}) = +0.096\ \mathrm{kg \cdot m/s}$. So the x component of the change in momentum is $\Delta p_{\mathrm{Ec}x} = (+0.096\ \mathrm{kg \cdot m/s}) - (+0.048\ \mathrm{kg \cdot m/s}) = +0.048\ \mathrm{kg \cdot m/s}$. ✔

(*b*) In reference frame M, $p_{\mathrm{Mc}x,\mathrm{i}} = mv_{\mathrm{Mc}x,\mathrm{i}} = (0.12\ \mathrm{kg}) \times (-0.20\ \mathrm{m/s}) = -0.024\ \mathrm{kg \cdot m/s}$ and $p_{\mathrm{Mc}x,\mathrm{f}} = mv_{\mathrm{Mc}x,\mathrm{f}} = (0.12\ \mathrm{kg})(+0.20\ \mathrm{m/s}) = +0.024\ \mathrm{kg \cdot m/s}$. So $\Delta p_{\mathrm{Mc}x} = (+0.024\ \mathrm{kg \cdot m/s}) - (-0.024\ \mathrm{kg \cdot m/s}) = +0.048\ \mathrm{kg \cdot m/s}$. ✔

4 EVALUATE RESULT The momentum change is the same in the two reference frames, as I expect. In reference frame M, the $v_x(t)$ curve shifts up or down, but its shape does not change, and therefore the values of $\Delta\vec{v}$ and $\Delta\vec{p}$ must also remain the same.

Example 6.3 reaffirms what we have already seen:

The change in a system's momentum is the same in any inertial reference frame.

6.4 In Example 6.3, what is the change in the cart's kinetic energy due to the shove (*a*) in the Earth reference frame, (*b*) in a reference frame moving in the same direction as the cart at 0.60 m/s relative to Earth, and (*c*) in a reference frame moving in the same direction as the cart at 0.80 m/s relative to Earth?

Unlike changes in momentum, changes in kinetic energy are not the same in two reference frames moving relative to each other. As Checkpoint 6.4 shows, the cart's kinetic energy increases, decreases, or stays the same depending on the reference frame.

The observation that changes in kinetic energy depend on the reference frame is disturbing. What does this result imply about conservation of energy? To begin addressing this question, let's look first at an elastic collision, such as the one in Figure 6.8. In **Table 6.1** I have listed the velocities and kinetic energy of both carts. Note that even though the change in velocity Δv_x for each cart is the same in the two reference frames, the change in kinetic energy ΔK for each cart is not. In a given reference frame, however, the absolute value of ΔK is the same for both carts, and so the change in the kinetic energy K of the two-cart system is zero in each frame.

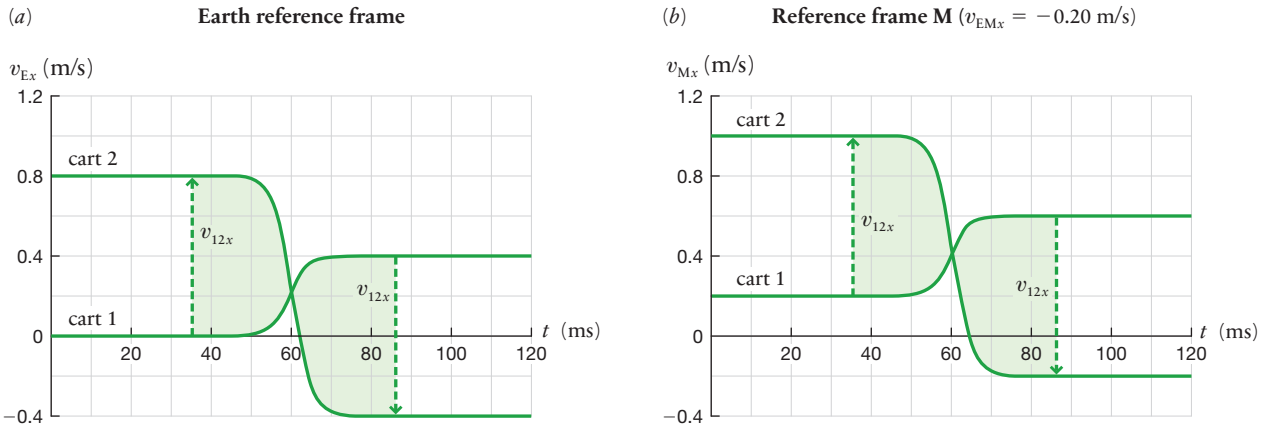
A simple way of arguing that the kinetic energy of the two-cart system *has* to remain unchanged in elastic

CONCEPTS

Table 6.1 An elastic collision seen from two reference frames

		v_x (m/s)				Kinetic energy (10^{-3} J)		
Cart	Inertia (kg)	*before*	*after*	Δv_x		*before*	*after*	ΔK
Earth reference frame								
1	0.36	0	+0.40	+0.40		0	29	+29
2	0.12	+0.80	−0.40	−1.2		38	9	−29
					K	38	38	0
Reference frame moving at −0.20 m/s relative to Earth								
1	0.36	+0.20	+0.60	+0.40		7	65	+58
2	0.12	+1.0	−0.20	−1.2		60	2	−58
					K	67	67	0

Figure 6.11 The relative velocity of the two colliding carts of Figure 6.8 is the same in (*a*) the Earth reference frame and (*b*) a reference frame moving along the track at $v_{EMx} = -0.20$ m/s relative to Earth.



collisions in any inertial reference frame is to look at how we formulated kinetic energy in Chapter 5 in terms of the relative velocities of two objects in a collision. As we saw earlier, differences in velocity do not depend on the velocity of the reference frame (Figure 6.11). In particular, our experimental observation that the relative speed of two colliding objects is unchanged in an elastic collision holds when we measure that relative speed in a reference frame that is moving at constant velocity. More generally, we observe:

The kinetic energy of a system of two elastically colliding objects does not change in any inertial reference frame.

Even though K does not change in elastic collisions, the fact that the individual ΔK values depend on the reference frame could be a big problem for conservation of energy in inelastic collisions. This reference-frame dependence for the individual ΔK values could mean that ΔK for the system also depends on the reference frame. Recall from Section 5.3 that in inelastic collisions some kinetic energy is converted to internal energy, which is a measure of the state of the system. Because changes in kinetic energy are reference-frame-dependent, the amount that gets converted could be reference-frame-dependent as well and that would mean that the final state of the system would depend on the reference frame! To determine whether or not this is the case, we examine an inelastic collision in the next example.

CONCEPTS

EXAMPLE 6.4 Energy conversion

Consider a collision between the two carts of Table 6.1, starting from the same initial velocities, but with $v_{E1x,f} = +0.30$ m/s. Make a table like Table 6.1 for this situation, and compare the amount of kinetic energy converted to internal energy in the Earth reference frame and in a reference frame M moving at -0.20 m/s relative to the track.

1 GETTING STARTED Because all the initial values are the same, I can copy all the "before" values from Table 6.1. I am given the final velocity of cart 1, and so I need to determine the final velocity of cart 2 in the Earth reference frame, then determine the final velocities in reference frame M, calculate the corresponding kinetic energies, and determine the change in kinetic energy in each reference frame.

2 DEVISE PLAN To determine $v_{E2x,f}$, I apply conservation of momentum to the isolated system of the two carts, which tells me that $\Delta\vec{p} = \vec{0}$. To calculate the two final velocities in reference frame M, I subtract the velocity of this reference frame from the carts' velocities in the Earth reference frame. Then I determine the kinetic energies from $K = \frac{1}{2}mv^2$ and complete the table by determining the changes in kinetic energy.

3 EXECUTE PLAN Because $\vec{p}_i = \vec{p}_f$, I have

$$m_1 v_{E1x,i} + m_2 v_{E2x,i} = m_1 v_{E1x,f} + m_2 v_{E2x,f}.$$

Substituting $v_{E1x,i} = 0$ and solving for $v_{E2x,f}$, I get

$$v_{E2x,f} = \frac{m_2 v_{E2x,i} - m_1 v_{E1x,f}}{m_2}.$$

Substituting the given values into this expression, I obtain $v_{E2x,f} = -0.10$ m/s.

Because $v_{EMx} = -0.20$ m/s, I must add 0.20 m/s to the x components of the carts' velocities in the Earth reference frame to determine the corresponding values in reference frame M:

$$v_{M1x,i} = 0 + (+0.20\text{ m/s}) = +0.20\text{ m/s}$$

$$v_{M2x,i} = (+0.80\text{ m/s}) + (+0.20\text{ m/s}) = +1.00\text{ m/s}$$

$$v_{M1x,f} = (+0.30\text{ m/s}) + (+0.20\text{ m/s}) = +0.50\text{ m/s}$$

$$v_{M2x,f} = (-0.10\text{ m/s}) + (+0.20\text{ m/s}) = +0.10\text{ m/s}.$$

Now I can calculate all the kinetic energies and complete the table (Table 6.2). ✔

Table 6.2 An inelastic collision seen from two reference frames

Cart	Inertia (kg)	v_x (m/s) before	v_x (m/s) after	Δv_x	Kinetic energy (10^{-3} J) before	Kinetic energy (10^{-3} J) after	ΔK
Earth reference frame							
1	0.36	0	+0.30	+0.30	0	16	+16
2	0.12	+0.80	−0.10	−0.90	38	1	−37
				K	38	17	−21
Reference frame moving at −0.20 m/s relative to Earth							
1	0.36	+0.20	+0.50	+0.30	7	45	+38
2	0.12	+1.0	+0.10	−0.90	60	1	−59
				K	67	46	−21

4 **EVALUATE RESULT** Each cart's kinetic energy change is different in the two reference frames because, as just noted in the text, changes in kinetic energy are reference-frame-dependent. However, the change in the *system's* kinetic energy is the same in the two reference frames: −21 mJ in both cases. This tells me that the amount of kinetic energy converted to internal energy—and therefore the final state of the two carts—does not depend on the reference frame, as I expect.



6.5 Repeat Example 6.4 but let the collision be totally inelastic.

The results of Example 6.4 and Checkpoint 6.5 are reassuring. Imagine that the converted kinetic energy becomes thermal energy, with the result that the temperature of each cart increases. You wouldn't expect these temperature changes to be different when you view the collision from reference frames that are moving at constant speed relative to each other. Likewise, in a collision between two cars, where the converted kinetic energy goes into crushing the fenders, you wouldn't expect the bending of the metal to depend on the reference frame. The amount of kinetic energy converted to internal energy during an inelastic collision does not depend on the velocity of the reference frame relative to the colliding objects.

These observations, together with the results we obtained in the examples in this section, tell us that the same laws and principles we discussed in Chapters 4 and 5 hold in any inertial reference frame. In particular, the momentum and energy laws can be used in any inertial reference frame. Earlier we saw that it is not possible to tell the difference between various inertial reference frames—that is, reference frames in which the law of inertia holds. It is an experimental fact that things behave the same way in all inertial reference frames. We should therefore expect the laws that describe that behavior to be the same too. This leads us to formulate the **principle of relativity:**

The laws of the universe are the same in all inertial reference frames, which move at constant velocity relative to each other.

The principle of relativity provides a criterion for judging theories: For a theory to be valid, it must prescribe the same behavior in all inertial reference frames. Although observers in different inertial reference frames record different values for the velocities of carts involved in a collision and therefore different momenta and kinetic energies, they agree that the momentum and energy of the two-cart system do not change.

It follows from the principle of relativity that no experiment carried out entirely in one inertial reference frame can tell what the motion of that reference frame is relative to another reference frame. Returning to the example of Section 6.2: Sitting in a very quiet airplane during a smooth, turbulence-free ride, it is impossible to tell if you are flying or if the airplane is sitting on the runway with its engines running. Of course, you can tell you are moving relative to Earth's surface by peeking out of the window and observing the landscape move by. By timing the motion of some reference points on Earth, you could even determine the velocity of the airplane relative to Earth. By peeking out of the window, however, your observations are no longer confined to a single reference frame. To determine your velocity relative to Earth's surface, you need to take measurements in two reference frames.

A consequence of the principle of relativity is that it is not possible to deduce from measurements taken entirely in one reference frame the motion of that reference frame relative to other reference frames.

This statement also means that no inertial reference frame is preferred over any other; it is not possible to determine the *absolute* velocity of any reference frame or object. The laws of the universe are the same in all inertial reference frames, and there is no reference frame that is "at rest" in some absolute sense.

CONCEPTS

When you are standing on Earth's surface, are you "at rest" or moving at 30 km/s (the speed of Earth in its orbit around the Sun), or is your speed higher still—say, the speed of the Sun in its galactic orbit? None of these statements is meaningful because no experiments can be done to determine your speed in empty space. You can only speak of your velocity or speed *relative to something else*.

When specifying the velocity of an object, we must therefore always state relative to what we measure that velocity. For simplicity, when the Earth reference frame is the only reference frame of interest, we will omit this relative statement and the subscript E. The phrase "his velocity was $\vec{v}$" will be understood to mean the velocity is measured relative to Earth.

6.6 Is the coefficient of restitution e different in two inertial reference frames, which are moving at constant velocity relative to each other? (See Eq. 5.18 if you have forgotten the definition of e.)

6.4 Zero-momentum reference frame

Let me end this first part of the chapter by addressing a practical question. If you can choose whichever inertial reference frame you wish, is any choice better than another? The reference frame you are in (usually one at rest relative to Earth) is a logical choice. By choosing some other reference frame, however, you could adjust the value of the momentum of a system up or down by some fixed amount without violating the conservation laws. You could, for instance, adjust the velocity of your reference frame in such a way that the momentum of the system you are observing becomes zero. Such a reference frame is called the system's **zero-momentum reference frame.** To distinguish the zero-momentum reference frame from other reference frames, we'll refer to quantities measured in that reference frame with the capital letter Z. As we shall see, examining collisions from this reference frame often simplifies things.

The prescription for calculating the velocity of the zero-momentum reference frame is as follows:

The velocity $\vec{v}_{EZ}$ of a system's zero-momentum reference frame relative to Earth is equal to the momentum of the system in the Earth reference frame divided by the inertia of the system.

EXAMPLE 6.5 Zeroing out the momentum

For the colliding carts in Figure 6.8, (*a*) determine the velocity of the zero-momentum reference frame relative to Earth and (*b*) show that the system momentum measured in the zero-momentum reference frame is zero both before and after the collision.

1 GETTING STARTED This example requires me to apply the prescription for calculating the velocity of the zero-momentum reference frame and then verify that the system momentum is indeed zero in that reference frame.

2 DEVISE PLAN To determine the velocity of the zero-momentum reference frame relative to Earth, I calculate the momentum of the two-cart system in the Earth reference frame and then divide this result by the sum of the inertias of the two carts.

Once I have determined the velocity of the zero-momentum reference frame, I subtract it from the carts' velocities in the Earth reference frame to obtain their velocities in the zero-momentum reference frame. I then use these velocities to calculate the carts' initial and final momenta.

3 EXECUTE PLAN (*a*) In the Earth reference frame, cart 1 (inertia 0.36 kg) is initially at rest and cart 2 (inertia 0.12 kg) is moving at +0.80 m/s. The x component of the combined momentum of the two carts is therefore $(0.12\ \text{kg})(+0.80\ \text{m/s}) = +0.096\ \text{kg}\cdot\text{m/s}$. The x component of the velocity of the zero-momentum reference frame is thus

$$\frac{0.096\ \text{kg}\cdot\text{m/s}}{0.36\ \text{kg} + 0.12\ \text{kg}} = +0.20\ \text{m/s}. \checkmark$$

(*b*) To obtain the x component of the velocities of the carts in the zero-momentum reference frame, I must subtract 0.20 m/s from their values in the Earth reference frame: $v_{Z1x,i} = -0.20\ \text{m/s}$ and $v_{Z2x,i} = +0.60\ \text{m/s}$. Then $p_{Z1x,i} = -0.072\ \text{kg}\cdot\text{m/s}$ and $p_{Z2x,i} = +0.072\ \text{kg}\cdot\text{m/s}$, making the initial momentum of the system measured from this reference frame zero. ✔

For the x component of the final velocities in the zero-momentum reference frame, I get $v_{Z1x,f} = +0.20\ \text{m/s}$ and $v_{Z2x,f} = -0.60\ \text{m/s}$. So $p_{Z1x,f} = +0.072\ \text{kg}\cdot\text{m/s}$ and $p_{Z2x,f} = -0.072\ \text{kg}\cdot\text{m/s}$, making the system's final momentum measured from this reference frame zero. ✔

4 EVALUATE RESULT In the zero-momentum reference frame, the momentum of the two-cart system before the collision is zero, as it must be from the definition of the zero-momentum frame. I also observe that it is zero after the collision, as I expect given that the system is isolated.

Figure 6.12 and Table 6.3 summarize the results of Example 6.5. Note how the shift from the Earth reference frame to the zero-momentum reference frame increases symmetry in the momentum-versus-time graph. In the zero-momentum reference frame, the velocity and momentum of each cart simply change sign in the collision. In this frame, the system as a whole is at rest, allowing us to concentrate on the motion of the objects within the system.

6.7 Is the kinetic energy of the two-cart system in Figure 6.12 in the zero-momentum reference frame less than, equal to, or greater than the system's kinetic energy in the Earth reference frame?

CONCEPTS

Figure 6.12 Collision between two carts as seen (*a*) from the Earth reference frame and (*b*) from the zero-momentum reference frame. The inertias of the carts are $m_1 = 0.36$ kg and $m_2 = 0.12$ kg and the x axis points to the right in the diagrams.

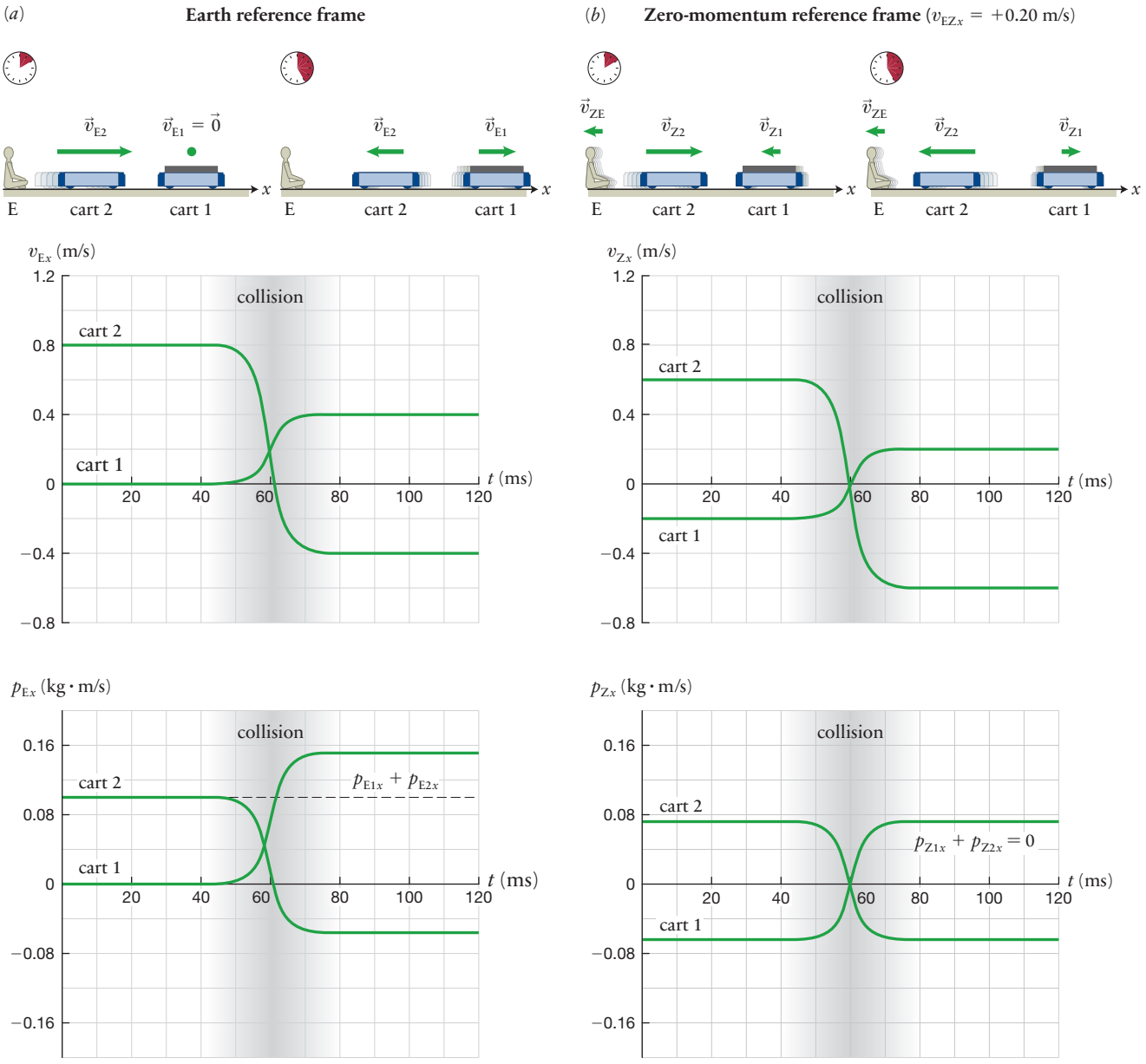


Table 6.3 Symmetry in a zero-momentum reference frame

	Velocity change (m/s)		Momentum change (kg · m/s)	
Cart	Earth reference frame	Zero-momentum reference frame	Earth reference frame	Zero-momentum reference frame
1	$0 \rightarrow +0.40$	$-0.20 \rightarrow +0.20$	$0 \rightarrow +0.14$	$-0.072 \rightarrow +0.072$
2	$+0.80 \rightarrow -0.40$	$+0.60 \rightarrow -0.60$	$+0.096 \rightarrow -0.048$	$+0.072 \rightarrow -0.072$

Self-quiz

1. A space traveler discovers an object that accelerates in her reference frame. Which conclusion is correct? (*a*) Her reference frame is noninertial. (*b*) The object is not isolated. (*c*) You cannot tell.
2. A jogger starts from rest along a straight track. Consider the jogger-Earth system to be isolated. As the jogger's speed increases, does the speed of Earth change or remain constant?
3. A driver slams on the brakes of a car in which a marble is glued to the dashboard. To an observer standing by the side of the road, the marble slows down and has a nonzero acceleration. This happens because (*a*) the observer is in a noninertial reference frame, (*b*) the marble is not isolated, (*c*) some other reason.
4. Is it always possible to choose a zero-momentum reference frame for an isolated system that contains more than two objects?
5. On a low-friction track, two carts with unequal inertias move toward each other. After the collision, the two stick together and are at rest relative to the track. Before the collision, a zero-momentum reference frame would (*a*) move in the direction of the cart that has the greater inertia, (*b*) move in the direction of the faster cart, (*c*) not move relative to the track. (*d*) Can't tell—you need to know the carts' inertias and velocities.

Answers

1. (*c*). To claim that her reference frame is inertial, you need to know whether or not an *isolated* object remains at rest in that reference frame. From the information given, you don't know whether the object is accelerating because it is not isolated or because the reference frame is noninertial.
2. It changes. Strange as it sounds, Earth's velocity does change every time you start moving! Otherwise, conservation of momentum would not hold. Can you verify this statement experimentally? Not for the jogger-Earth case, because the ratio $\Delta v_{\text{jogger}}/\Delta v_{\text{Earth}}$ is equal to the inverse ratio of the inertias. Because Earth's inertia is enormous, its velocity change is infinitesimally small. This being so, how can you be sure you have the correct answer? Certainty comes from knowing that the law of conservation of momentum has been found to hold everywhere it has ever been tested—including on a cosmic scale.
3. (*b*). From the point of view of the observer, if the marble were not glued to the dashboard, the inertia of the marble would keep it moving at the velocity it had before the driver hit the brakes. The marble slows down only because the car's dashboard holds it in place. In other words, the dashboard and the marble interact with each other, and therefore the marble is not isolated.
4. Yes. The momentum of any system of objects is the sum of the individual momenta. To obtain the momentum of an object of inertia m_o in a reference frame moving at velocity $\vec{v}_{\text{EM}}$ relative to Earth, we must subtract $m_o\vec{v}_{\text{EM}}$ from the momentum of that object in the Earth reference frame. The system's momentum in that reference frame is then equal to the momentum of the system in the Earth reference frame minus the product $(m_1 + m_2 + m_3 + \cdots)\vec{v}_{\text{EM}}$. We can always choose a velocity $\vec{v}_{\text{EM}}$ such that the magnitude of this product is equal to the magnitude of the momentum of the system in the Earth reference frame and the algebraic signs are opposites.
5. (*c*). After the collision, both carts come to rest in the reference frame of the track, which means that the momentum of the two-cart system after the collision is zero. This means that the system's momentum before the collision was also zero. Therefore the reference frame of the track is a zero-momentum reference frame.

6.5 Galilean relativity

Consider two observers, A and B, moving at constant velocity relative to each other. Suppose they observe the same *event*—such as a firecracker exploding—and describe it relative to their respective reference frames and clocks (**Figure 6.13**). Let the origins of the two observers' reference frames coincide at $t = 0$ (Figure 6.13*a*). Observer A sees the event as happening at position $\vec{r}_{Ae}$ at clock reading t_{Ae} (Figure 6.13*b*).* Observer B sees the event at position $\vec{r}_{Be}$ at clock reading t_{Be}. What is the relationship between these clock readings and positions?

If, as we discussed in Chapter 1, we assume time is absolute—the same everywhere—and if the two observers have synchronized their (identical) clocks, they both observe the event at the same clock readings, which means

$$t_{Ae} = t_{Be}. \tag{6.1}$$

Because the clock readings of the two observers always agree, we can omit the subscripts referring to the reference frames:

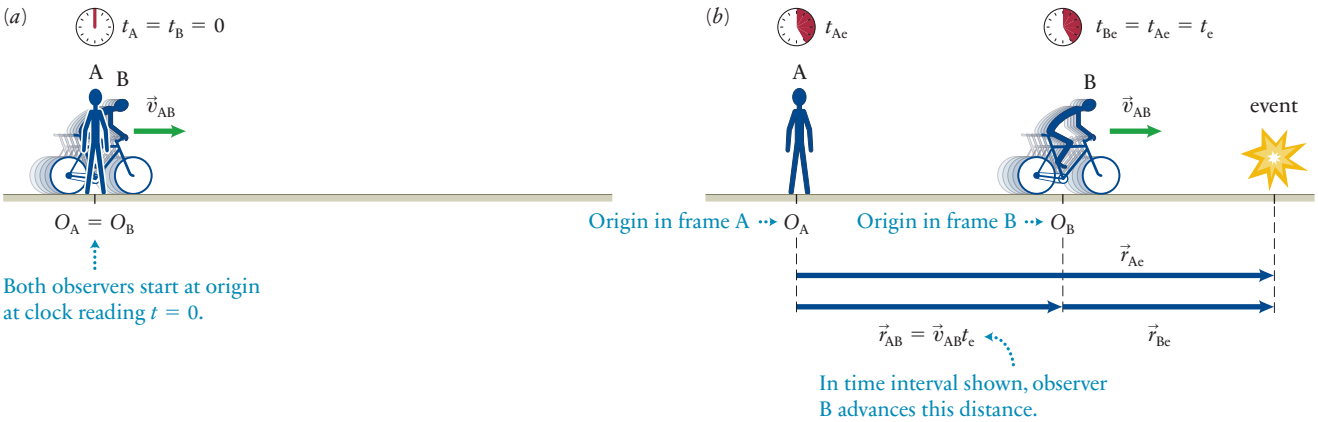
$$t_A = t_B = t. \tag{6.2}$$

From Figure 6.13 we see that the position $\vec{r}_{AB}$ of observer B in reference frame A at instant t_e is equal to B's displacement over the time interval $\Delta t = t_e - 0 = t_e$, and so $\vec{r}_{AB} = \vec{v}_{AB}t_e$ because B moves at constant velocity $\vec{v}_{AB}$. Therefore

$$\vec{r}_{Ae} = \vec{r}_{AB} + \vec{r}_{Be} = \vec{v}_{AB}t_e + \vec{r}_{Be}. \tag{6.3}$$

Equations 6.2 and 6.3 allow us to relate event data collected in one reference frame to data on the same event e collected in a reference frame that moves at constant velocity relative to the first one (neither of these has to be at rest relative to Earth, but their origins must coincide at $t = 0$). To this end we

Figure 6.13 Two observers moving relative to each other observe the same event. Observer B moves at constant velocity $\vec{v}_{AB}$ relative to observer A. (*a*) The origins *O* of the two reference frames overlap at instant $t = 0$. (*b*) At instant t_e, when the event occurs, the origin of observer B's reference frame has a displacement $\vec{v}_{AB}t_e$ relative to reference frame A.



*Remember our subscript form: The capital letter refers to the reference frame; the lowercase e is for "event." Thus the vector $\vec{r}_{Ae}$ represents observer A's measurement of the position at which the event occurs.

rewrite these equations so that they give the values of time and position in reference frame B in terms of quantities measured in reference frame A:

$$t_B = t_A = t \tag{6.4}$$

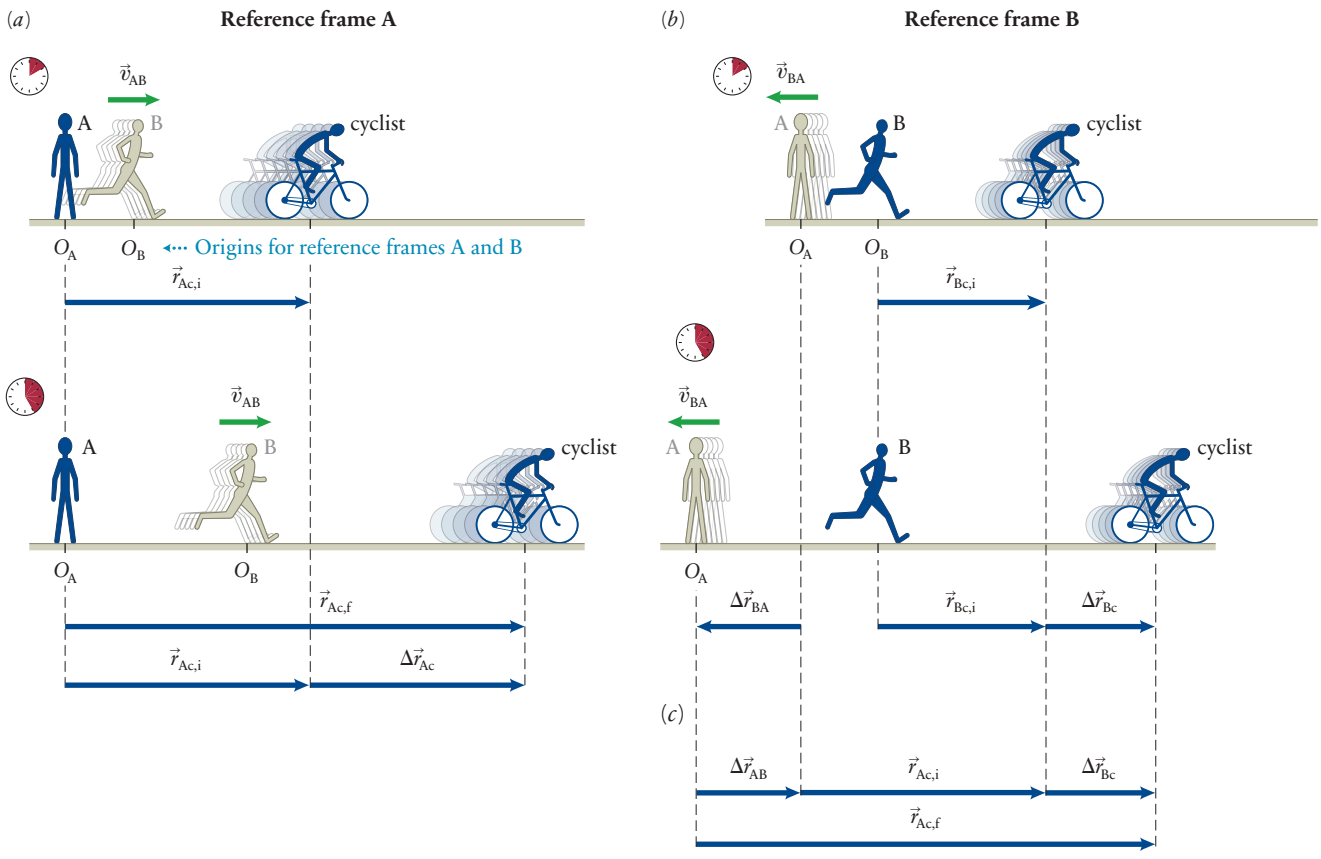
$$\vec{r}_{Be} = \vec{r}_{Ae} - \vec{v}_{AB}t_e. \tag{6.5}$$

These two equations, called the **Galilean transformation equations,** form the basis for *Galilean relativity*; they allow us to relate observations made in different reference frames. The Galilean transformation equations are predicated on the assumption that measurements of time intervals and lengths are not affected by motion. Even though this assumption sounds reasonable, it has been invalidated by experimental confirmation of the theory of special relativity (see Chapter 14). For most everyday phenomena, however, when speeds are much lower than the speed of light (about 3.0×10^8 m/s, see Section 1.3), Eqs. 6.4 and 6.5 remain excellent approximations.

6.8 A jogger runs in place on a treadmill whose belt moves at $v_{EBx} = +2.0$ m/s relative to Earth. Let the origins of the Earth reference frame and the reference frame B moving along with the top surface of the belt coincide at $t = 0$. (*a*) What is the jogger's position in the reference frame of the belt at $t = 10$ s? (*b*) Use Eq. 6.3 to show that an Earth observer's measurement of the jogger's position is $r_{Ejx} = 0$ at all instants.

Now that we have determined the relationship between positions in two reference frames, let us examine the relationship between the velocities of a cyclist relative to two observers, A and B, who are moving at constant velocity $\vec{v}_{AB}$ relative to each other (**Figure 6.14**). Each observer

Figure 6.14 The cyclist's displacement $\Delta\vec{r}_{Ac}$ measured by observer A is different from the displacement $\Delta\vec{r}_{Bc}$ measured by observer B, who is moving relative to observer A.



determines the velocity of the cyclist by measuring the cyclist's displacement during a certain time interval Δt.

According to observer A, the bicycle undergoes a displacement $\Delta\vec{r}_{Ac}$ in the time interval Δt (Figure 6.14*a*). The bicycle's displacement $\Delta\vec{r}_{Bc}$ measured by observer B, however, is different because B undergoes a displacement $\Delta\vec{r}_{AB}$ relative to A (Figure 6.14*b*). Comparing the displacements, we obtain (Figure 6.14*c*)

$$\vec{r}_{Ac,f} - \vec{r}_{Ac,i} = \Delta\vec{r}_{Ac} = \Delta\vec{r}_{AB} + \Delta\vec{r}_{Bc}. \tag{6.6}$$

If we divide the two sides of this equation by the time interval Δt, we obtain

$$\frac{\Delta\vec{r}_{Ac}}{\Delta t} = \frac{\Delta\vec{r}_{AB}}{\Delta t} + \frac{\Delta\vec{r}_{Bc}}{\Delta t}, \tag{6.7}$$

and if we let Δt approach zero, we get

$$\vec{v}_{Ac} = \vec{v}_{AB} + \vec{v}_{Bc}. \tag{6.8}$$

Rewriting this equation so that it gives the velocity of an object o in reference frame B in terms of quantities measured in reference frame A, we get

$$\vec{v}_{Bo} = \vec{v}_{Ao} - \vec{v}_{AB}. \tag{6.9}$$

Equation 6.8 tells us that velocities obtained in different reference frames are additive. If an object moves at a velocity $\vec{v}_{Bo}$ in a reference frame that is moving at a velocity $\vec{v}_{AB}$ relative to some observer A, then the velocity $\vec{v}_{Ao}$ of that object in A's reference frame is the sum $\vec{v}_{AB} + \vec{v}_{Bo}$.

As an example, imagine walking in a train that moves along a straight track at 25 m/s relative to Earth. If you are walking forward at 1.0 m/s relative to the train, what is your velocity relative to Earth? In 1.0 s, the train moves 25 m relative to Earth, and in that same time interval you move that same 25 m plus an additional 1.0 m. Because you move (25 m) + (1.0 m) = 26 m each second, your velocity relative to Earth is 26 m/s.

To determine the relationship between an object's acceleration in the two reference frames, we begin by writing what a change in the object's velocity in reference frame B corresponds to in reference frame A. Using Eq. 6.8 and keeping in mind that the relative velocity $\vec{v}_{AB}$ of the two reference frames is constant, we get

$$\begin{aligned}\Delta\vec{v}_{Ao} &= \vec{v}_{Ao,f} - \vec{v}_{Ao,i} = (\vec{v}_{AB} + \vec{v}_{Bo,f}) - (\vec{v}_{AB} + \vec{v}_{Bo,i}) \\ &= \vec{v}_{Bo,f} - \vec{v}_{Bo,i} = \Delta\vec{v}_{Bo}.\end{aligned} \tag{6.10}$$

In words, *changes in velocity are the same in any two reference frames moving at constant relative velocity*—a conclusion we reached in Section 6.3.

Dividing both sides of Eq. 6.10 by Δt and taking the limit as Δt approaches zero, we determine that the accelerations in the two reference frames are identical, too:

$$\vec{a}_{Ao} \equiv \lim_{\Delta t \to 0} \frac{\Delta\vec{v}_{Ao}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta\vec{v}_{Bo}}{\Delta t} \equiv \vec{a}_{Bo}. \tag{6.11}$$

In particular, if an object moves at constant velocity in the Earth reference frame ($\vec{a}_{Eo} = \vec{0}$), then it moves at constant velocity ($\vec{a}_{Mo} = \vec{0}$) in any reference frame M that moves at constant velocity relative to Earth, confirming what we concluded in Section 6.1.

Two bullet trains passing each other. If each train has a speed of 50 m/s relative to the track, at what speed does the oncoming train approach the one from which the photo is taken?



Answer: 100 m/s.

QUANTITATIVE TOOLS

To simplify working with quantities in various reference frames, note from Figure 6.13*b* that

$$\vec{r}_{Ae} = \vec{r}_{AB} + \vec{r}_{Be} \tag{6.12}$$

The first and last subscripts on either side of this equation are the same, as indicated by the arrows. Moreover, if you imagine canceling the two identical subscripts B that follow each other on the right side and contracting the two vectors $\vec{r}$ into one, you get

$$\vec{r}_{AB} + \vec{r}_{Be} = \vec{r}_{Ae}, \tag{6.13}$$

which is the same as Eq. 6.12. This "subscript cancellation" works with any relative quantity. For example, observer A's measurement of the velocity of an object is equal to observer A's measurement of the velocity of observer B plus observer B's measurement of the velocity of the object:

$$\vec{v}_{Ao} = \vec{v}_{AB} + \vec{v}_{Bo}, \tag{6.14}$$

which is the same as Eq. 6.8.

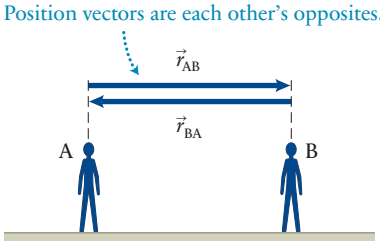
Figure 6.15 shows another useful subscript operation: Reversing the order of the subscripts reverses the direction of a relative vector $\vec{r}_{AB}$. The vector describing the position of observer B relative to A has the same magnitude as the vector $\vec{r}_{BA}$ describing the position of observer A relative to B, but it points in the opposite direction:

$$\vec{r}_{AB} = -\vec{r}_{BA}, \tag{6.15}$$

and by taking the derivative with respect to time of this equation, we obtain a similar expression for the velocity:

$$\vec{v}_{AB} = -\vec{v}_{BA}. \tag{6.16}$$

Figure 6.15 The position vectors of two observers relative to each other are of equal magnitude and point in opposite directions.



PROCEDURE: Applying Galilean relativity

In problems dealing with more than one reference frame, you need to keep track not only of objects, but also of reference frames. For this reason, each quantity is labeled with two subscripts. The first subscript denotes the observer; the second denotes the object of interest. For example, if we have an observer on a train and also a car somewhere on the ground but in sight of the train, then $\vec{a}_{Tc}$ is the train observer's measurement of the acceleration of the car. Once you understand this notation and a few basic operations, working with relative quantities is easy and straightforward.

observer A's measurement of velocity of car:

$$\vec{v}_{Ac}$$

observer ..... object of interest

1. Begin by listing the quantities given in the problem, using this double-subscript notation.
2. Write the quantities you need to determine in the same notation.
3. Use subscript cancellation (Eq. 6.13) to write an equation for each quantity you need to determine, keeping the first and the last subscripts on each side the same. For example, in a problem where you need to determine $\vec{v}_{Tc}$ involving a moving observer B, write

$$\vec{v}_{Tc} = \vec{v}_{TB} + \vec{v}_{Bc}.$$

4. If needed, use subscript reversal (Eq. 6.15) to eliminate any unknowns.
5. Use the kinematics relationships from Chapters 2 and 3 to solve for any remaining unknowns, making sure you stay in one reference frame.

You can use this procedure and the subscript operations for any of the three basic kinematic quantities (position, velocity, and acceleration).

QUANTITATIVE TOOLS

EXAMPLE 6.6 Who is moving?

You are driving at 25 m/s on a straight, horizontal road when a truck going 30 m/s in the same direction overtakes you. Let the positive x direction point in the direction of travel, and let the origins of the reference frames affixed to your car and the truck coincide at the instant the truck overtakes you. (*a*) What is your car's velocity as measured by someone in the truck? (*b*) What is the velocity of the truck relative to your car? (*c*) What is your car's position as measured by someone in the truck 60 s after overtaking you?

❶ **GETTING STARTED** This example involves two reference frames: the Earth reference frame (E) and the reference frame moving along with the truck (T). The velocities of the car and the truck are given in the Earth reference frame: $v_{\mathrm{Ec}x} = +25$ m/s and $v_{\mathrm{ET}x} = +30$ m/s. I need to determine the velocity of my car $\vec{v}_{\mathrm{Tc}}$ relative to the truck, the velocity of the truck $\vec{v}_{\mathrm{cT}}$ relative to my car, and my car's position $\vec{r}_{\mathrm{Tc}}$ relative to the truck 60 s after the overtaking.

❷ **DEVISE PLAN** I begin by using subscript cancellation to write an equation that gives the velocity $\vec{v}_{\mathrm{Tc}}$ of my car relative to the truck in terms of velocities involving the Earth reference frame: $\vec{v}_{\mathrm{Tc}} = \vec{v}_{\mathrm{TE}} + \vec{v}_{\mathrm{Ec}}$. The value of $\vec{v}_{\mathrm{Ec}}$ is given (+25 m/s), but I don't have a value for $\vec{v}_{\mathrm{TE}}$. However, I can use subscript reversal to write $v_{\mathrm{TE}x} = -v_{\mathrm{ET}x} = -30$ m/s. Once I have $\vec{v}_{\mathrm{Tc}}$, I can use subscript reversal once more to determine $\vec{v}_{\mathrm{cT}}$. The product of $\vec{v}_{\mathrm{Tc}}$ and the 60-s time interval gives me $\vec{r}_{\mathrm{Tc}}$ (from Eq. 2.16, $\Delta x = v_x \Delta t$).

❸ **EXECUTE PLAN** (*a*) Using subscript cancellation and reversal, I get $\vec{v}_{\mathrm{Tc}} = \vec{v}_{\mathrm{TE}} + \vec{v}_{\mathrm{Ec}} = -\vec{v}_{\mathrm{ET}} + \vec{v}_{\mathrm{Ec}}$ or, in terms of x components,

$$v_{\mathrm{Tc}x} = -v_{\mathrm{ET}x} + v_{\mathrm{Ec}x} = -(+30\ \mathrm{m/s}) + (+25\ \mathrm{m/s})$$

$$= -5.0\ \mathrm{m/s}. \checkmark$$

(*b*) $v_{\mathrm{cT}x} = -v_{\mathrm{Tc}x} = +5.0$ m/s ✔

(*c*) My displacement in the truck's reference frame is $\Delta x_{\mathrm{Tc}} = (-5.0\ \mathrm{m/s})(60\ \mathrm{s}) = -300$ m. Thus 60 s after the truck passes you, a person in the truck measures your position as $r_{\mathrm{Tc}x} = 300$ m behind the truck. ✔

❹ **EVALUATE RESULT** The truck is moving faster than my car, and so relative to the truck my car moves in the negative x direction, confirming the minus sign on my answers to parts *a* and *c*. Relative to me, the truck moves in the positive x direction, in agreement with the plus sign of my answer to part *b*.



6.9 In a train moving due north at 3.1 m/s relative to Earth, a passenger carrying a suitcase walks due north down the aisle at 1.2 m/s relative to the train. A spider crawls along the bottom of the suitcase at 0.5 m/s due south relative to the suitcase. What is the velocity of the spider relative to Earth?

6.6 Center of mass

To relate momentum in two reference frames moving relative to each other, we must examine inertia. Equation 4.2 told us how to obtain the inertia m_{o} of an object by letting the object collide with the standard inertia:

$$m_{\mathrm{o}} \equiv -\frac{\Delta v_{\mathrm{s}x}}{\Delta v_{\mathrm{o}x}} m_{\mathrm{s}}. \tag{6.17}$$

Suppose this inertia m_{o} is measured by two observers, A and B, who are moving at constant velocity $\vec{v}_{\mathrm{AB}}$ relative to each other. Even though the velocity values measured by the two observers differ, Eq. 6.10 tells us that the moving observer obtains the same velocity changes $\Delta v_{\mathrm{s}x}$ and $\Delta v_{\mathrm{o}x}$ as those measured in the Earth reference frame. This means that we can omit the subscript referring to the reference frame and write

$$m_{\mathrm{Ao}} = m_{\mathrm{Bo}} = m_{\mathrm{o}}. \tag{6.18}$$

Therefore the momenta of an object measured in two reference frames A and B are related by

$$\vec{p}_{\mathrm{Ao}} \equiv m_{\mathrm{o}} \vec{v}_{\mathrm{Ao}} = m_{\mathrm{o}}(\vec{v}_{\mathrm{AB}} + \vec{v}_{\mathrm{Bo}}) \equiv m_{\mathrm{o}} \vec{v}_{\mathrm{AB}} + \vec{p}_{\mathrm{Bo}}, \tag{6.19}$$

where we have used Eq. 6.8 to relate the velocities in the two reference frames. Equation 6.19 shows that the momentum measured in reference

frame B is different from the momentum measured in reference frame A. In particular, if the velocity of reference frame B relative to reference frame A is $\vec{v}_{AB} = \vec{p}_{Ao}/m_o$, the momentum measured in reference frame B is zero: $\vec{p}_{Bo} = \vec{0}$ (and so reference frame B is a zero-momentum reference frame for that object).

We can make the momentum of a system of objects zero in the same way. Suppose we have a system made up of objects that have inertias $m_1, m_2, \ldots$, and we let the momentum of the system in reference frame A be $\vec{p}_{A\,sys}$. Using Eq. 6.19, we can write for the momentum of the system

$$\begin{aligned}\vec{p}_{A\,sys} &= \vec{p}_{A1} + \vec{p}_{A2} + \cdots \\ &= (m_1\vec{v}_{AB} + \vec{p}_{B1}) + (m_2\vec{v}_{AB} + \vec{p}_{B2}) + \cdots \\ &= (m_1 + m_2 + \cdots)\vec{v}_{AB} + (\vec{p}_{B1} + \vec{p}_{B2} + \cdots) = m\vec{v}_{AB} + \vec{p}_{B\,sys}, \quad (6.20)\end{aligned}$$

where $m = m_1 + m_2 + \cdots$ is the inertia of the system. If we adjust the velocity $\vec{v}_{AB}$ of reference frame B relative to reference frame A such that

$$\vec{v}_{AB} = \frac{\vec{p}_{A\,sys}}{m}, \quad (6.21)$$

then, according to Eq. 6.20, reference frame B is a zero-momentum reference frame:

$$\vec{p}_{B\,sys} = \vec{p}_{A\,sys} - m\vec{v}_{AB} = \vec{p}_{A\,sys} - m\frac{\vec{p}_{A\,sys}}{m} = \vec{0}. \quad (6.22)$$

Equations 6.21 and 6.22 show that relative to Earth the velocity of the zero-momentum reference frame Z for a system of objects is equal to the system's momentum measured in the Earth reference frame divided by the inertia of the system (as we stated in Section 6.4):

$$\vec{v}_{EZ} = \frac{\vec{p}_{E\,sys}}{m} = \frac{m_1\vec{v}_{E1} + m_2\vec{v}_{E2} + \cdots}{m_1 + m_2 + \cdots} \quad \begin{matrix}\text{(zero-momentum} \\ \text{reference frame)}\end{matrix}. \quad (6.23)$$

The velocity of the zero-momentum reference frame is related to the position of the **center of mass** of a system. This position is defined as

$$\vec{r}_{cm} \equiv \frac{m_1\vec{r}_1 + m_2\vec{r}_2 + \cdots}{m_1 + m_2 + \cdots}, \quad (6.24)$$

where $\vec{r}_1, \vec{r}_2, \ldots$ represent the positions of the objects of inertia $m_1, m_2, \ldots$ in the system.* Equation 6.24 applies to the values of the positions measured in *any* reference frame (and for this reason we can omit the reference-frame subscripts). To show the relationship between the velocity of the zero-momentum reference frame and the position of the center of mass, we take the derivative of the left and right sides of Eq. 6.24 with respect to time:

$$\frac{d\vec{r}_{cm}}{dt} = \frac{m_1(d\vec{r}_1/dt) + m_2(d\vec{r}_2/dt) + \cdots}{m_1 + m_2 + \cdots}. \quad (6.25)$$

*Even though I pointed out in Chapter 4 that *mass* and *inertia* are related and they are essentially synonyms, we should—to be consistent with our use of the word *inertia* for the quantity m—call $\vec{r}_{cm}$ the position of the *center of inertia*. Historically, however, $\vec{r}_{cm}$ has always been called the position of the *center of mass*, a tradition we shall follow.

Because $d\vec{r}/dt = \vec{v}$ (Eq. 2.25), we can rewrite Eq. 6.25 in the form

$$\vec{v}_{\text{cm}} \equiv \frac{d\vec{r}_{\text{cm}}}{dt} = \frac{m_1\vec{v}_1 + m_2\vec{v}_2 + \cdots}{m_1 + m_2 + \cdots}, \tag{6.26}$$

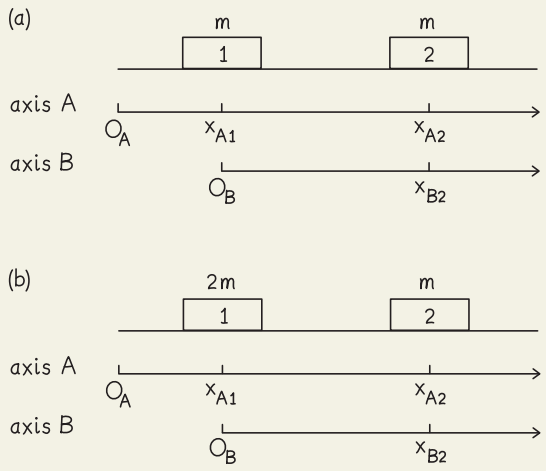
where $\vec{v}_{\text{cm}}$ is the system's **center-of-mass velocity.** Note that this velocity is precisely the velocity of the zero-momentum reference frame in Eq. 6.23. Thus we conclude that, when measured from a reference frame moving at the same velocity as the center of mass of a system, the system's momentum is automatically zero. (In other words, such a reference frame is a zero-momentum reference frame.) In figures we shall denote the center of mass by a circle with a cross in it: ⊗.

EXAMPLE 6.7 Two-cart center of mass

The positions of two identical carts at rest on a low-friction track are measured relative to two axes oriented in the same direction along the track. On axis A, cart 1 is at x_{A1} and cart 2 is at x_{A2}. On axis B, cart 1 is at the origin and cart 2 is at x_{B2}. (*a*) Determine the position of the center of mass of the two-cart system relative to each axis. (*b*) If instead of identical carts you have $m_1 = 2m_2$, what is the position of the center of mass on each axis?

❶ **GETTING STARTED** I begin by making a sketch of the situation (**Figure 6.16**). Each axis corresponds to a reference frame that is at rest relative to Earth. Because the origins of the two axes don't coincide, the x components of the cart positions on axis A are different from those on axis B. I have to use these x components to determine the position of the system's center of mass on each axis for two cases: when the cart inertias are identical and when they are not. (For simplicity, the carts are treated as points; the positions given are the positions of these points.)

Figure 6.16



❷ **DEVISE PLAN** The position of the center of mass of a system of objects is given by Eq. 6.24. To use this equation in this problem, I must write it in terms of components. For axis A, I have

$$x_{\text{Acm}} = \frac{m_1x_{\text{A1}} + m_2x_{\text{A2}}}{m_1 + m_2}, \tag{1}$$

and for axis B I replace the subscripts A with B. To solve this problem, I substitute the given values into this equation.

❸ **EXECUTE PLAN** (*a*) Substituting $m_1 = m_2 = m$ into Eq. 1, I get for axis A

$$x_{\text{Acm}} = \frac{mx_{\text{A1}} + mx_{\text{A2}}}{m + m} = \frac{m(x_{\text{A1}} + x_{\text{A2}})}{2m} = \frac{x_{\text{A1}} + x_{\text{A2}}}{2}. \checkmark$$

This position is halfway between the the two carts (see Figure 6.16). For axis B, I get

$$x_{\text{Bcm}} = \frac{mx_{\text{B1}} + mx_{\text{B2}}}{m + m} = \frac{m(0 + x_{\text{B2}})}{2m} = \frac{x_{\text{B2}}}{2}, \checkmark$$

which is again the position midway between the two carts.

(*b*) Now I have $m_1 = 2m_2 = 2m$, and so for axis A,

$$x_{\text{Acm}} = \frac{2mx_{\text{A1}} + mx_{\text{A2}}}{2m + m} = \frac{2x_{\text{A1}} + x_{\text{A2}}}{3}. \checkmark$$

This position is one-third of the way from x_{A1} to x_{A2} on axis A because $(2x_{\text{A1}} + x_{\text{A2}})/3 = x_{\text{A1}} + (x_{\text{A2}} - x_{\text{A1}})/3$. For axis B,

$$x_{\text{Bcm}} = \frac{2m \cdot (0) + mx_{\text{B2}}}{2m + m} = \frac{x_{\text{B2}}}{3}, \checkmark$$

which is one-third of the way between the origin of axis B and cart 2. If you go back to the carts in Figure 6.16*b* and mark the center-of-mass position on both axes, you'll see that the position $x_{\text{Bcm}} = x_{\text{B2}}/3$ is aligned with the position $x_{\text{Acm}} = (2x_{\text{A1}} + x_{\text{A2}})/3$, as it must be. The center of mass of any given system can be at one position only, regardless of which reference frame is used to locate that position.

❹ **EVALUATE RESULT** When $m_1 = m_2$, the center of mass lies at the center of the system (midway between the carts) regardless of the choice of axis, which makes sense. When $m_1 = 2m_2$, the center of mass lies twice as close to cart 1 as to cart 2. This makes sense because I expect the center of mass to shift toward the object with the larger inertia. If, for example, the inertia of cart 1 were so large that the inertia of cart 2 was negligible, I could drop the second term in the numerator and denominator of Eq. 1, which would make the position of the center of mass be that of cart 1.

6.10 In Example 6.7, let $m_1 = 3m_2$. (*a*) Where on axis A is the center of mass of the two-cart system? (*b*) Where on axis A would you need to place a third cart of inertia $m_3 = m_1$ so that the center of mass of the three-cart system is at the position of cart 2?

Example 6.7 demonstrates an important point: The position of the center of mass of a system is a property of the system that is independent of the choice of reference frame. Equation 6.24 suggests that the center of mass represents a sort of "average" position of a system, but the center of mass is more important than that. Back in Chapter 2 we began our study of motion by specifying the position of an object at each instant by the position vector $\vec{r}$. Real objects, however, are *extended* rather than *point-like,* and so to say "the position" of a system or of a real object is not a precise statement unless we specify a fixed reference point in the system or on the object. For example, the statement "I parked my car 3 m from the fire hydrant" could mean that the front end of the car is 3 m away from the hydrant, or the midpoint of the car is 3 m from the hydrant, or any other point on the car is 3 m from the hydrant. The center of mass allows us to specify a fixed position in a system according to an exact prescription given by Eq. 6.24.

For extended objects, we can apply Eq. 6.24 by dividing the object into small pieces and substituting the position and inertia of each piece into the equation. This method is described in Appendix D. For symmetrical objects that have a uniformly distributed inertia, however, we don't need to carry out the calculation because the center of mass lies on an axis of symmetry and on any plane of symmetry. For example, the center of mass of a uniform rod lies at the center of the rod, midway between its ends, and the center of mass of a uniform sphere or cube lies at the geometric center of the sphere or cube.

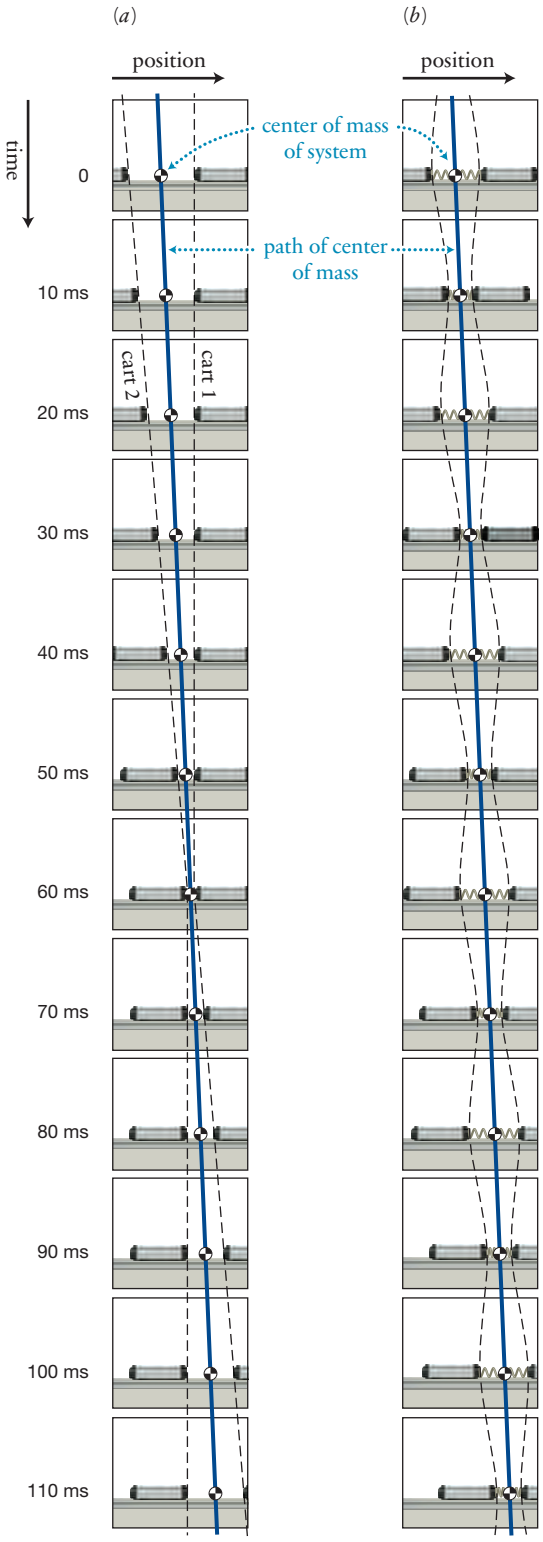
Center of mass is also an important tool in our quest for simplification: It allows us to be precise even for systems where the motion is complex. To take a specific example, consider the motions shown in **Figure 6.17**. Because the two carts are identical, the center of mass of the two-cart system is always halfway between them. Notice that the center-of-mass motion is unaffected by the collision in Figure 6.17*a*. This is so because the momentum of the system doesn't change, and so the motion of the system must be constant. The center of mass of the system moves at velocity $\vec{v}_{\text{cm}}$ both before and after the collision.

The importance of the center of mass becomes even more evident in Figure 6.17*b*. Now the carts bounce back and forth because of the action of the spring while the two-cart system as a whole moves to the right. The motion of the center of mass of the system is entirely independent of the motion of the carts relative to each other and independent of the interactions between the carts. Even though the system continuously changes shape (the distance between the carts keeps changing), the center of mass provides a fixed reference point that moves at constant velocity.

6.11 (*a*) Determine the center-of-mass velocity of the two carts in Figure 6.8*a* before and after the collision, and verify that it is equal to the velocity of the carts at the point where the two $v_x(t)$ curves intersect. (*b*) Is the velocity of the carts at the point where the two $v_x(t)$ curves intersect always equal to the center-of-mass velocity? (*c*) Determine the velocities of the two carts in the zero-momentum reference frame, before and after the collision, and show that in this reference frame the momentum of the system is always zero.

QUANTITATIVE TOOLS

Figure 6.17 Two identical carts moving on a low-friction track. (*a*) The carts collide elastically. (*b*) The carts are connected by a spring and bounce back and forth about the center of mass. In both cases, the center of mass moves at constant velocity, so that the $x(t)$ curve for the center of mass is always a straight line.



6.7 Convertible kinetic energy

We can derive some additional useful expressions involving the center-of-mass velocity. Keeping in mind that $\vec{v}_{\text{Eo}} = \vec{v}_{\text{EZ}} + \vec{v}_{\text{Zo}}$ (Eq. 6.8) and that for the zero-momentum reference frame $\vec{v}_{\text{EZ}} = \vec{v}_{\text{cm}}$ (Eqs. 6.23 and 6.26), we can derive an expression that gives, for a system of objects, the kinetic energy $K_{\text{E sys}}$ measured in the Earth reference frame in terms of the corresponding kinetic energy $K_{\text{Z sys}}$ measured in the zero-momentum reference frame:

$$\begin{aligned} K_{\text{E sys}} &= \tfrac{1}{2}m_1 v_{\text{E1}x}^2 + \tfrac{1}{2}m_2 v_{\text{E2}x}^2 + \cdots \\ &= \tfrac{1}{2}m_1 (v_{\text{cm}x} + v_{\text{Z1}x})^2 + \tfrac{1}{2}m_2 (v_{\text{cm}x} + v_{\text{Z2}x})^2 + \cdots . \end{aligned} \tag{6.27}$$

Working out the first term on the right gives

$$\tfrac{1}{2}m_1 (v_{\text{cm}x} + v_{\text{Z1}x})^2 = \tfrac{1}{2}m_1 v_{\text{cm}}^2 + m_1 v_{\text{cm}x} v_{\text{Z1}x} + \tfrac{1}{2}m_1 v_{\text{Z1}}^2, \tag{6.28}$$

where I dropped the subscripts x on the terms that contain squares of the x component of the velocity because $v_x^2 = v^2$. Working out the other terms in Eq. 6.27 and gathering terms that contain equal powers of $v_{\text{cm}x}$, we obtain

$$\begin{aligned} K_{\text{E sys}} &= \tfrac{1}{2}(m_1 + m_2 + \cdots)v_{\text{cm}}^2 + (m_1 v_{\text{Z1}x} + m_2 v_{\text{Z2}x} + \cdots)v_{\text{cm}x} \\ &\quad + \left(\tfrac{1}{2}m_1 v_{\text{Z1}}^2 + \tfrac{1}{2}m_2 v_{\text{Z2}}^2 + \cdots\right). \end{aligned} \tag{6.29}$$

From the definition of the zero-momentum reference frame, we have

$$m_1 v_{\text{Z1}x} + m_2 v_{\text{Z2}x} + \cdots = p_{\text{Z1}x} + p_{\text{Z2}x} + \cdots = (p_{\text{Z sys}})_x = 0, \tag{6.30}$$

so that the middle term on the right in Eq. 6.29 disappears. The last term in Eq. 6.29 is the kinetic energy $K_{\text{Z sys}}$ of the system measured in the zero-momentum reference frame, and so substituting $m = m_1 + m_2 + \cdots$, we have

$$K_{\text{E sys}} = \tfrac{1}{2}mv_{\text{cm}}^2 + K_{\text{Z sys}}. \tag{6.31}$$

The first term on the right in this equation, called the system's **translational kinetic energy,** is the kinetic energy associated with the motion of the center of mass of the system:

$$K_{\text{cm}} \equiv \tfrac{1}{2}mv_{\text{cm}}^2. \tag{6.32}$$

This term has the form of the kinetic energy of an object of inertia m moving at speed v_{cm}.

For an isolated system, the translational kinetic energy K_{cm} is *nonconvertible*, which means that it cannot be converted to internal energy. To see why this is true, note that this kinetic energy is a function of the system's center-of-mass velocity, which cannot change for an isolated system (see Section 6.8).

The remainder of the kinetic energy—the $K_{\text{Z sys}}$ term in Eq. 6.31—is the system's **convertible kinetic energy,** the amount that can be converted to internal energy without changing the momentum of the system. It is equal to the system's kinetic energy minus the (nonconvertible) translational kinetic energy:

$$\begin{aligned} K_{\text{conv}} \equiv K_{\text{Z sys}} &= K_{\text{E sys}} - \tfrac{1}{2}mv_{\text{cm}}^2 \\ &= \left(\tfrac{1}{2}m_1 v_{\text{E1}}^2 + \tfrac{1}{2}m_2 v_{\text{E2}}^2 + \cdots\right) - \tfrac{1}{2}(m_1 + m_2 + \cdots)v_{\text{cm}}^2. \end{aligned} \tag{6.33}$$

With this last equation, we can eliminate any reference to the zero-momentum reference frame in Eq. 6.31, leaving only quantities measured in the Earth reference frame:

$$K_{\text{E sys}} = K_{\text{cm}} + K_{\text{conv}}. \tag{6.34}$$

If we had started not from an expression for the kinetic energy in the Earth reference frame (Eq. 6.27) but from the corresponding expression in any other inertial reference frame A, we would have obtained the same results, with A substituted for all the subscripts E. For this reason, we can omit the subscript E and rewrite Eqs. 6.31–6.34 in a form that holds in any inertial reference frame:

$$K = K_{\text{cm}} + K_{\text{conv}}, \tag{6.35}$$

where
$$K_{\text{conv}} = \left(\tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2 + \cdots\right) - \tfrac{1}{2}mv_{\text{cm}}^2. \tag{6.36}$$

The kinetic energy of a system can be split into a convertible part and a nonconvertible part. The nonconvertible part is the system's translational kinetic energy $K_{\text{cm}} = \frac{1}{2}mv_{\text{cm}}^2$. The remainder of the kinetic energy is convertible.

For a system of two colliding objects, the convertible kinetic energy is

$$K_{\text{conv}} = K - \tfrac{1}{2}mv_{\text{cm}}^2 = \left(\tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2\right) - \tfrac{1}{2}(m_1 + m_2)v_{\text{cm}}^2. \tag{6.37}$$

With some algebra (see Checkpoint 6.12), substituting the Eq. 6.26 form of v_{cm} into this expression yields

$$K_{\text{conv}} = \tfrac{1}{2}\left(\frac{m_1m_2}{m_1 + m_2}\right)v_{12}^2 \quad \text{(two-object system)}. \tag{6.38}$$

If I now write

$$\mu \equiv \frac{m_1m_2}{m_1 + m_2}, \tag{6.39}$$

then the convertible kinetic energy for two objects takes the simple form

$$K_{\text{conv}} = \tfrac{1}{2}\mu v_{12}^2 \qquad \text{(two-object system)}, \tag{6.40}$$

where $\vec{v}_{12} = \vec{v}_2 - \vec{v}_1$ is the relative velocity of the two objects (see Eq. 5.1). The right side of this expression has the form of the kinetic energy of an object of inertia μ moving at a speed v_{12}. The quantity represented by the Greek letter μ (mu), which we introduced to simplify our notation, thus has the same units as inertia. It is called the **reduced inertia,** or *reduced mass*, because it is less than the inertia of either of the two colliding objects.



6.12 Verify that Eq. 6.38 is valid by substituting Eq. 6.26 for v_{cm} into Eq. 6.37 and working through the algebra.

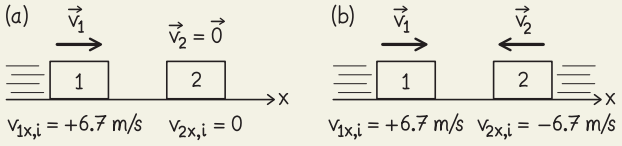
EXAMPLE 6.8 Crash test

Two collisions are carried out to crash-test a 1000-kg car. (*a*) While moving at 15 mi/h, the car strikes an identical car initially at rest. (*b*) While moving at 15 mi/h, the car strikes an identical car moving toward it and also traveling at 15 mi/h. For each collision, how much kinetic energy can be converted to internal energy, and what fraction of the initial kinetic energy of the two-car system does this represent?

❶ **GETTING STARTED** I am given the inertias of two cars and their initial velocities in two collisions. I must calculate two values for each collision: the amount of kinetic energy converted during the collision and what fraction of the system's energy this converted amount is. I begin by making a sketch of the situation before each collision (**Figure 6.18**). I assume that the car that is initially at rest is free to roll (its parking brake is not on), and I ignore any friction so that I can consider the two-car system to be isolated. I also convert the speed from miles per hour to meters per second:

$$\frac{15 \text{ miles}}{1 \text{ hour}} \times \frac{1 \text{ hour}}{3600 \text{ s}} \times \frac{1609 \text{ m}}{1 \text{ mile}} = 6.7 \text{ m/s}.$$

Figure 6.18



❷ **DEVISE PLAN** The amount of energy that can be converted is given by Eq. 6.40, which requires me to calculate the reduced inertia of the two-car system (Eq. 6.39) and the relative speed of the two cars (Eq. 5.1). To determine the fraction of the initial kinetic energy that this converted amount of energy represents, I divide the converted amount by the sum of the initial kinetic energies $\frac{1}{2}mv^2$ of the two cars.

❸ **EXECUTE PLAN** (*a*) Substituting $m_1 = m_2 = 1000$ kg into Eq. 6.39, I obtain the reduced inertia of the two-car system: $\mu = 500$ kg. The initial relative speed of the two cars is

$$v_{12} = |v_{2x} - v_{1x}| = |0 - (+6.7 \text{ m/s})| = 6.7 \text{ m/s}$$

and so

$$K_{\text{conv}} = \tfrac{1}{2}(500 \text{ kg})(6.7 \text{ m/s})^2 = 11 \text{ kJ}. \checkmark$$

The initial kinetic energy of the system is $K_i = K_{1,i} = \frac{1}{2}(1000 \text{ kg})(6.7 \text{ m/s})^2 = 22$ kJ, and so the fraction of energy converted is $K_{\text{conv}}/K_i = (11 \text{ kJ})/(22 \text{ kJ}) = 0.50$. ✔

(*b*) The initial relative speed of the two cars is

$$v_{12} = |v_{2x} - v_{1x}| = |6.7 \text{ m/s} - (-6.7 \text{ m/s})| = 13.4 \text{ m/s}$$

and so $K_{\text{conv}} = \frac{1}{2}(500 \text{ kg})(13.4 \text{ m/s})^2 = 45$ kJ. ✔

The initial kinetic energy of the system is $K_i = K_{1,i} + K_{2,i} = \frac{1}{2}(1000 \text{ kg})(6.7 \text{ m/s})^2 + \frac{1}{2}(1000 \text{ kg})(6.7 \text{ m/s})^2 = 45$ kJ, and so all of the initial kinetic energy is convertible: $K_{\text{conv}}/K_i = (45 \text{ kJ})/(45 \text{ kJ}) = 1.0$. ✔

❹ **EVALUATE RESULT** When both cars move at 15 mi/h, the system has twice as much kinetic energy as when only one car moves at that speed, which makes sense. When only one car moves, however, the momentum of the system is nonzero, which means that not all the kinetic energy can be converted, as I found. In the second collision, the momentum of the system is zero, and so all of the kinetic energy can be converted, in agreement with my answer.

For an inelastic collision, v_{12} changes and so K changes, too. Using Eq. 6.35, we can write for the change in the kinetic energy of the system:

$$\Delta K = \Delta K_{\text{cm}} + \Delta K_{\text{conv}}. \tag{6.41}$$

The first term on the right is zero because the system is isolated, and so its translational kinetic energy cannot change. Substituting Eq. 6.40 into the last term, we have

$$\Delta K = \tfrac{1}{2}\mu v_{12f}^2 - \tfrac{1}{2}\mu v_{12i}^2 = \tfrac{1}{2}\mu(v_{12f}^2 - v_{12i}^2). \tag{6.42}$$

If we pull v_{12i}^2 out of the parentheses, we can write this expression in terms of the coefficient of restitution e (see Eq. 5.18):

$$\Delta K = \tfrac{1}{2}\mu v_{12i}^2\left(\frac{v_{12f}^2}{v_{12i}^2} - 1\right) = \tfrac{1}{2}\mu v_{12i}^2(e^2 - 1). \tag{6.43}$$

This value represents the amount of kinetic energy converted to internal energy during the inelastic collision. The maximum change in the system's kinetic energy occurs in a totally inelastic collision ($e = 0$):

$$\Delta K = -\tfrac{1}{2}\mu v_{12i}^2 \qquad \text{(totally inelastic collision)}. \tag{6.44}$$

So, when two colliding objects stick together, their relative velocity becomes zero and all of the convertible kinetic energy is converted to internal energy (compare Eqs. 6.40 and 6.44). Even though the collision is totally inelastic, however, the two-object system still has some kinetic energy because the nonconvertible part of the system's kinetic energy ($\frac{1}{2}mv_{\text{cm}}^2$) is unaffected by the collision.

For an elastic collision, $e = 1$, and so Eq. 6.43 becomes $\Delta K = 0$: The kinetic energy is the same before and after the collision, as expected.

6.13 A moving object that has inertia m strikes a stationary object that has inertia $0.5m$. (*a*) What fraction of the initial kinetic energy of the system is convertible? (*b*) Why can't the rest be converted?

6.8 Conservation laws and relativity

In Section 6.3 we introduced the principle of relativity, which states that the laws of the universe are the same in any inertial reference frame. In this section we prove this statement for the momentum and energy laws, which embody the laws of conservation of momentum and energy.

Let us consider a system from two reference frames, A and B, that are moving at velocity $\vec{v}_{\text{AB}}$ relative to each other. In Section 6.6 we saw that the momentum of an object is different in two reference frames that are moving relative to each other (Eq. 6.19), but now we are interested in *changes* in momentum. For an object of inertia m_{o}, we can write

$$\Delta\vec{p}_{\text{Ao}} = m_{\text{o}}\Delta\vec{v}_{\text{Ao}}. \tag{6.45}$$

Because a change in velocity in reference frame A is the same as a change in reference frame B (Eq. 6.10), we have

$$\Delta\vec{p}_{\text{Ao}} = m_{\text{o}}\Delta\vec{v}_{\text{Ao}} = m_{\text{o}}\Delta\vec{v}_{\text{Bo}} = \Delta\vec{p}_{\text{Bo}}, \tag{6.46}$$

which tells us that the change in momentum of an object is the same in reference frames A and B. So, for the momentum of a system of objects, we have

$$\Delta\vec{p}_{\text{A sys}} = \Delta\vec{p}_{\text{B sys}}. \tag{6.47}$$

In words, *changes in the momentum of a system are the same in any two reference frames moving at constant velocity relative to each other.*

Equation 6.47 tells us that if the change in momentum of a system is zero in reference frame A, then it is also zero in reference frame B. So, if a system is isolated in reference frame A, it is also isolated in reference frame B, as required by the principle of relativity (**Figure 6.19**).

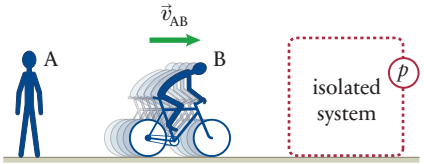
Next we examine changes in the energy of a system. Let us first consider a change in the system's internal energy. In Section 5.3 we introduced internal energy as a quantitative measure of a change in state. Given that the state of any object or system cannot depend on the motion of the observer, we must conclude that any change in internal energy must be independent of the reference frame:

$$\Delta E_{\text{A int}} = \Delta E_{\text{B int}}. \tag{6.48}$$

Unfortunately, the situation is not so simple for a change in kinetic energy, as we saw in Section 6.3. If an object's velocity in reference frame A increases to $v_{\text{Ao}x,\text{i}} + \Delta v_{\text{Ao}x}$ from an initial value $v_{\text{Ao}x,\text{i}}$, its kinetic energy increases by an amount

$$\begin{aligned}\Delta K_{\text{Ao}} &= \tfrac{1}{2}m_{\text{o}}[(v_{\text{Ao}x,\text{i}} + \Delta v_{\text{Ao}x})^2 - (v_{\text{Ao}x,\text{i}})^2] \\ &= m_{\text{o}}v_{\text{Ao}x,\text{i}}\Delta v_{\text{Ao}x} + \tfrac{1}{2}m_{\text{o}}(\Delta v_{\text{Ao}x})^2.\end{aligned} \tag{6.49}$$

Figure 6.19 Whether or not a system is isolated does not depend on the motion of the observer, as long as the reference frames are inertial relative to each other.



Substituting Eqs. 6.8 and 6.10 into the right side, we get

$$\Delta K_{\mathrm{Ao}} = m_{\mathrm{o}}(v_{\mathrm{AB}x} + v_{\mathrm{Bo}x,\mathrm{i}})\Delta v_{\mathrm{Bo}x} + \tfrac{1}{2}m_{\mathrm{o}}(\Delta v_{\mathrm{Bo}x})^2. \tag{6.50}$$

Applying Eq. 6.49 to reference frame B, we have

$$\Delta K_{\mathrm{Bo}} = m_{\mathrm{o}}v_{\mathrm{Bo}x,\mathrm{i}}\Delta v_{\mathrm{Bo}x} + \tfrac{1}{2}m_{\mathrm{o}}(\Delta v_{\mathrm{Bo}x})^2. \tag{6.51}$$

When we expand Eq. 6.50, we get

$$\Delta K_{\mathrm{Ao}} = m_{\mathrm{o}}v_{\mathrm{AB}x}\Delta v_{\mathrm{Bo}x} + m_{\mathrm{o}}v_{\mathrm{Bo}x,\mathrm{i}}\Delta v_{\mathrm{Bo}x} + \tfrac{1}{2}m_{\mathrm{o}}(\Delta v_{\mathrm{Bo}x})^2. \tag{6.52}$$

Because the second and third terms on the right are exactly what we have on the right in Eq. 6.51, we see that

$$\Delta K_{\mathrm{Ao}} = m_{\mathrm{o}}v_{\mathrm{AB}x}\Delta v_{\mathrm{Bo}x} + \Delta K_{\mathrm{Bo}} \neq \Delta K_{\mathrm{Bo}} \qquad (\text{single object}), \tag{6.53}$$

which means that the change in an object's kinetic energy depends on the reference frame in which that change is measured.

To investigate whether or not ΔK, the change in the kinetic energy of a system made up of two colliding objects, is independent of the reference frame in which the energy is measured, let us examine the kinetic energy converted to internal energy, given by Eq. 6.42. In reference frame A we have

$$\Delta K_{\mathrm{A}} = \tfrac{1}{2}\mu(v_{A12,\mathrm{f}}^2 - v_{A12,\mathrm{i}}^2) \qquad (\text{two-object system}). \tag{6.54}$$

Because the relative velocities are *differences* in velocities, Eq. 6.10 tells us that $v_{\mathrm{B12}} = v_{\mathrm{A12}}$, and so the change in the system's kinetic energy must also be the same in any two inertial reference frames:

$$\Delta K_{\mathrm{B}} = \Delta K_{\mathrm{A}} \tag{6.55}$$

(in spite of the fact that the changes in kinetic energy are different for the individual objects!). Combining Eqs. 6.55 and 6.48 and generalizing the result to a system of more than two objects, we obtain

$$\Delta K_{\mathrm{B}} + \Delta E_{\mathrm{B\,int}} = \Delta K_{\mathrm{A}} + \Delta E_{\mathrm{A\,int}} \tag{6.56}$$

or

$$\Delta E_{\mathrm{A\,sys}} = \Delta E_{\mathrm{B\,sys}}. \tag{6.57}$$

In words, *changes in the energy of a system are the same in any two reference frames moving at constant velocity relative to each other*.

Equation 6.57 tells us that if the change in energy of a system is zero in reference frame A, then it is also zero in reference frame B. So, if a system is closed in reference frame A, it is also closed in reference frame B, as required by the principle of relativity.

QUANTITATIVE TOOLS



6.14 Objects 1 ($m_1 = 1.0$ kg) and 2 ($m_2 = 3.0$ kg) collide inelastically. The velocities are $v_{1x,\mathrm{i}} = +4.0$ m/s, $v_{2x,\mathrm{i}} = 0$, $v_{1x,\mathrm{f}} = -0.50$ m/s, and $v_{2x,\mathrm{f}} = +1.5$ m/s. (*a*) What is the coefficient of restitution e? (*b*) Make a table like Table 6.1 showing the kinetic energy converted to internal energy both in the Earth reference frame and in a reference frame moving at $v_{\mathrm{EM}x} = -1.0$ m/s relative to Earth.

Chapter Summary

Inertial reference frames (Sections 6.1–6.4, 6.6)

Concepts A **reference frame** is a combination of a reference axis that defines a direction in space and a reference point that defines the origin from which motion is measured. The **Earth reference frame** is a reference frame at rest relative to Earth.

An **inertial reference frame** is one in which the **law of inertia** holds, which means a reference frame in which an isolated object remains either at rest or in motion at constant velocity.

The **principle of relativity** states that the laws of the universe are the same in all inertial reference frames.

The **zero-momentum reference frame** for a system of objects is the reference frame in which the momentum of the system is zero. The velocity of this reference frame is equal to the velocity of the center of mass of the system.

Quantitative Tools The position $\vec{r}_{cm}$ of the **center of mass** of a system of objects of inertias $m_1, m_2, \ldots$ located at positions $\vec{r}_1, \vec{r}_2, \ldots$ is

$$\vec{r}_{cm} \equiv \frac{m_1\vec{r}_1 + m_2\vec{r}_2 + \cdots}{m_1 + m_2 + \cdots}. \tag{6.24}$$

The **center-of-mass velocity** of a system of objects is

$$\vec{v}_{cm} \equiv \frac{d\vec{r}_{cm}}{dt} = \frac{m_1\vec{v}_1 + m_2\vec{v}_2 + \cdots}{m_1 + m_2 + \cdots}. \tag{6.26}$$

This is also the velocity of the **zero-momentum reference frame** for this system.

Relative motion (Section 6.5)

Concepts The **Galilean transformation equations** allow us to transform quantities measured in one inertial reference frame into their values in another inertial reference frame when the two reference frames are moving at constant relative velocity. The transformation equations tell us that *time intervals*, *lengths*, and *accelerations* are the same in any two inertial reference frames moving at constant relative velocity.

Quantitative Tools The **Galilean transformation equations** relate the time t and position $\vec{r}_e$ for an event (e) measured in an inertial reference frame A to these quantities measured for the event in any other inertial reference frame B. If the reference frames have a constant relative velocity $\vec{v}_{AB}$ and have their origins coinciding at $t = 0$, the transformation equations are

$$t_B = t_A = t \tag{6.4}$$

and

$$\vec{r}_{Be} = \vec{r}_{Ae} - \vec{v}_{AB}t_e. \tag{6.5}$$

As a consequence of these equations, the velocity $\vec{v}_{Ao}$ of an object (o) measured in an inertial reference frame A is related to the object's velocity measured in any other inertial reference frame B by

$$\vec{v}_{Ao} = \vec{v}_{AB} + \vec{v}_{Bo}. \tag{6.14}$$

The transformation equations also give the relationship between accelerations measured in any two inertial reference frames A and B:

$$\vec{a}_{Ao} \equiv \vec{a}_{Bo}. \tag{6.11}$$

Convertible kinetic energy (Section 6.7)

Concepts The **translational kinetic energy** of a system is the kinetic energy associated with the motion of its center of mass. For an isolated system, this kinetic energy is **nonconvertible** because it cannot be converted to internal energy (if it were, the momentum of the system would not be constant).

The **convertible kinetic energy** of an isolated system is the portion of the

Quantitative Tools The **translational (nonconvertible) kinetic energy** K_{cm} of a system is

$$K_{cm} \equiv \tfrac{1}{2}mv_{cm}^2, \tag{6.32}$$

where m is the system's inertia and v_{cm} is the speed of its center of mass.

The **convertible kinetic energy** of a two-particle system is

$$K_{conv} = \tfrac{1}{2}\mu v_{12}^2, \tag{6.40}$$

system's kinetic energy that can be converted to internal energy. This energy is the same in all inertial reference frames.

The **kinetic energy** of a system can be split into a convertible part K_{conv} and a nonconvertible part. The nonconvertible part is the system's translational kinetic energy K_{cm}.

where μ is the **reduced inertia** (or *reduced mass*) of the system, given by

$$\mu \equiv \frac{m_1 m_2}{m_1 + m_2}. \tag{6.39}$$

The kinetic energy of a system is

$$K = K_{cm} + K_{conv}. \tag{6.35}$$

Conservation laws and relativity (Section 6.8)

Concepts Changes in the momentum and energy of a system are the same in any two inertial reference frames.

Quantitative Tools If A and B are any two reference frames moving at constant velocity relative to each other, then the changes in the momentum and energy of any system are

$$\Delta \vec{p}_{A\,sys} = \Delta \vec{p}_{B\,sys} \tag{6.47}$$

$$\Delta K_B + \Delta E_{B\,int} = \Delta K_A + \Delta E_{A\,int}. \tag{6.56}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

6.1 Relativity of motion

1. Moving sidewalks are commonplace at large airports, but you must use caution getting on and off them. What should you do to make the transition easy when stepping onto one of these sidewalks? When stepping off? •
2. A flying bug hits the helmet of a coasting bicyclist. (What kind of collision do you expect this to be?) Draw the momentum vectors for the bug and the bicyclist before and after the collision (*a*) in the reference frame in which the bug is initially at rest and (*b*) in the reference frame in which the bicyclist is initially at rest. •
3. The data in the table were recorded for two race cars, A and B, driving down a straight stretch of track. Determine the average velocity of car B as measured by an observer in car A. ••

Time (s)	Car A position (m)	Car B position (m)
1	40	20
2	65	50
3	90	80
4	115	110
5	140	140
6	165	170
7	190	200
8	215	230
9	240	260
10	265	290

4. An observer in the Earth reference frame observes cart A and cart B moving toward each other with the same speed until they collide and come to rest. Consider this event as seen by observer A moving along with cart A. Assume that this observer matches cart A's initial velocity but does not collide with cart B (that is, observer A continues along at the same velocity that cart A had initially). Would the observer in the Earth reference frame and observer A obtain the same value for (*a*) the initial velocity of cart A, (*b*) the final velocity of cart B after the collision, and (*c*) the relative velocity of the two carts after the collision? ••
5. As you are standing on the platform, a boy in a train moving to the right extends his hand outside the window and releases a ball. In what direction is the ball moving according to (*a*) you, and (*b*) the boy? Ignore air resistance. ••
6. Two objects, A and B, of equal inertia approach each other with relative velocity $\vec{v}_{AB}$ and collide elastically. For each object, draw a velocity-versus-time graph for the interval starting a few seconds before the collision and ending a few seconds after the collision (*a*) from the reference frame in which A is initially at rest, (*b*) from the reference frame in which B is initially at rest, and (*c*) from the zero-momentum reference frame. ••

6.2 Inertial reference frames

7. You are watching a race while sitting on a bench. An athlete speeds up from 0 to 7 m/s to the right. Describe his motion from the point of view of another athlete going at (*a*) 5 m/s to the right, (*b*) 5 m/s to the left, and (*c*) accelerating from 0 to 5m/s to the right. •

8. On a long bus ride, you walk from your seat to the back of the bus to use the restroom. If the bus is driving at 100 km/h, and you walk at 2.0 m/s from your seat to the restroom, how quickly are you moving relative to the ground? •
9. A pickup truck has several empty soda cans loose in the bed. Why do the cans roll forward in the bed when the truck slows down? •
10. Does a man standing in an elevator that is moving upward at a constant speed of 10 m/s feel heavier or lighter? What if the elevator is moving downward at the same speed? If the elevator slows down to a stop? ••
11. Train A, $m_A = 150{,}000$ kg, is traveling west at 60 km/h. Train B, $m_B = 100{,}000$ kg, behind train A on the same track, is traveling west at 88 km/h and so is gaining on train A. Because the engineer driving B fails to slow down, B runs into the back of A. The two trains stick together and then move as a single unit after the collision. What is the momentum of each train before the collision and after the collision according to (*a*) an observer standing alongside the tracks and (*b*) an observer in a westbound automobile traveling at 100 km/h on a road that parallels the tracks? ••
12. A 1000-kg car traveling east at 50 km/h passes over the top of a hill and hits a 3000-kg truck stalled in the middle of the lane. The impact causes the truck to roll eastward at 15 km/h. (*a*) What is the coefficient of restitution for the collision? During the collision, how much kinetic energy is converted to internal energy (*b*) in the Earth reference frame and (*c*) in a reference frame moving at 5 km/h eastward along the road? ••
13. You are riding a 450-kg horse at 14.4 km/h east along a desert road. You have inertia equal to 60.0 kg. A police officer driving past (whom you know and who knows your inertia and the horse's inertia) measures your speed relative to the police car and calculates your kinetic energy to be 16.32 kJ. What possible speed(s) could the police car have been driving at the instant the officer measured your speed? •••

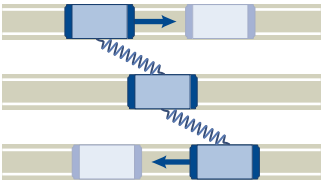
6.3 Principle of relativity

14. Two cars collide head-on on a busy street. An observer standing on the street witnesses the accident and calculates how much of the cars' initial kinetic energy went into deforming the cars upon collision, E_{def}. A police officer was driving next to one of the cars at a matching speed when the cars collided. Would this officer calculate a higher value of E_{def}, a lower value of E_{def}, or the same value of E_{def} if he used the velocities of the cars measured in his reference frame? •
15. A horizontal escalator connecting airport terminals is 1 km long. Two kids start from opposite ends and run toward each other at 3 m/s and meet 600 m from the start of the escalator. What is the speed of the escalator? How long do they take to reach each other? •
16. Do the laws of the universe change if the clocks in two inertial reference frames, A and B, are not synchronized (in other words, if $t_A = t_B + \Delta\tau$, where $\Delta\tau$ is some constant time interval)? ••
17. You place a 0.10-kg sonic ranger on a low-friction track in front of a 0.50-kg cart to measure the cart's velocity in the Earth reference frame, which turns out to be $+(1.0\ \text{m/s})\,\hat{\imath}$. You are distracted, the cart hits the ranger in a totally inelastic collision, and the two objects then move forward together. A friend is running toward the cart with a velocity of $-(3.0\ \text{m/s})\,\hat{\imath}$ with her sonic ranger on and pointed at the cart. (*a*) What is the momentum of the cart, with the ranger stuck to it, in your friend's reference frame? (*b*) What velocity does her sonic ranger measure for the cart after the collision? ••
18. At the roller rink, two 20-kg girls accelerate toward each other until they are each moving at 2.0 m/s in the Earth reference frame. They then collide stomach-to-stomach, grab on to each other, and fall to the floor. Calculate the momentum of each girl before the collision and after the collision (*a*) in the Earth reference frame and (*b*) in the reference frame of a mother who started skating with one of the girls and then continued on after the collision without changing speed. (*c*) The fathers decide the collision looks like fun and give it a try. Explain, in terms of energy, why they don't enjoy it nearly as much as their daughters do. ••
19. Carts A and B are identical and are moving toward each other on a track. The speed of cart A is v, while the speed of cart B is $2v$. In the Earth reference frame, the system of the two carts has kinetic energy K. Is there any other reference frame in which the two-cart system has the same kinetic energy K? If so, describe this reference frame. Otherwise, explain why no such reference frame exists. •••

6.4 Zero-momentum reference frame

20. Two identical cars approach each other head-on while traveling at the same speed. What does an observer in the Earth reference frame measure for the speed of the zero-momentum reference frame of this two-car system? Suppose the same cars are traveling in the same direction. Now what does an observer in the Earth reference frame measure for the speed of the zero-momentum reference frame? •
21. Two identical particles of inertia m collide elastically on a low-friction table. Calculate the kinetic energy of the system before and after the collision in the zero-momentum frames. •
22. Two identical particles, A and B, collide elastically. In the zero-momentum reference frame, what can you say about the ratios of the final and initial speeds, $v_{A,f}/v_{A,i}$ and $v_{B,f}/v_{B,i}$? •
23. Figure P6.23 shows three identical carts placed on rails (so that they can slide easily left and right) and connected to one another by springs. There are two simple, symmetrical ways for this system to vibrate in the zero-momentum reference frame. One way is the sweeping-X pattern shown in the figure. Draw the other way. ••

Figure P6.23



24. Two objects A and B having inertias $m_a = m$ and $m_b = 3m$ are moving with velocities $v_a = v$ and $v_b = 3v$. Find the velocities of the two objects in the zero-momentum frame of reference. ••
25. A 4000-kg dump truck is parked on a hill. The parking brake fails, and the truck rolls down the hill. It then coasts briefly along a horizontal stretch of road at 36 km/h before hitting a stationary 1000-kg car. The car sticks to the grill of the truck, and the two vehicles continue moving forward. (*a*) What is the speed of the combined unit immediately after the collision? (*b*) What is the momentum of the combined unit in the reference frame of a jogger approaching the collision at 2.0 m/s and moving in the direction opposite the direction of the velocity of the moving unit? (*c*) How fast, and in which direction, must the jogger be running in order for the momentum of the combined unit to be zero in his reference frame? ••
26. In the Earth reference frame, box 1 is approaching box 2, which is initially at rest on a low-friction floor, with velocity v. Box 1 has five times the inertia of box 2. They collide elastically. Calculate the velocities before and after collision in (*a*) the Earth reference frame and (*b*) the zero-momentum reference frame. •••
27. In an elastic collision between a lightweight object and a heavy object, which one carries away more of the kinetic energy? Does the answer depend on the initial speeds? (Hint: Begin with the zero-momentum reference frame for clarity.) •••

6.5 Galilean relativity

28. When two identical objects traveling at the same speed collide head-on, an observer standing in the Earth reference frame sees both objects changing direction. Do observers in every inertial reference frame that is moving relative to the Earth reference frame also see both objects changing direction? •
29. You toss a ball into the air and note the time interval between the ball leaving your hand and reaching its greatest height above you. While you are doing this, a construction worker being lifted on a hydraulic platform at constant speed also notes the time interval needed for the ball to reach its greatest height relative to himself. Is the time interval reported by the worker *longer, shorter,* or *the same as* the interval you report? ••
30. A student runs an experiment with two carts on a low-friction track. As measured in the Earth reference frame, cart 1 ($m = 0.36$ kg) moves from left to right at 1.0 m/s as the student walks along next to it at the same velocity. (*a*) What velocity $\vec{v}_{E2,i}$ in the Earth reference frame must cart 2 ($m = 0.12$ kg) have before the collision if, in the student's reference frame, cart 2 comes to rest right after the collision and cart 1 reverses direction and travels from right to left at 0.33 m/s? (*b*) What does the student measure for the momentum of the two-cart system? (*c*) What does a person standing in the Earth reference frame measure for the momentum of each cart before the collision? ••
31. Notice that in Figure 6.12 the time intervals of the interaction are all the same, regardless of the reference frame shown. Is this true in all reference frames? Why or why not? ••
32. In a three-car crash, car A bumps into the back end of car B, which then goes forward and bumps into the back end of car C. Is the distance that car B moves between the collisions the same in all inertial reference frames? ••
33. Riding up an escalator while staying on the same step for the whole ride takes 30 s. Walking up the same escalator takes 20 s. How long does it take to walk down the up escalator? ••
34. You run at 4 m/s in the same direction as a river flowing at 2 m/s. This speed is enough to keep you chatting with your friend who is steering a slow boat on the river. How fast is your friend going? How fast should your friend go, and in what direction, if the river suddenly speeds up to 3 m/s? ••
35. Airline pilots who fly round trips know that their round-trip travel time increases if there is any wind. To see this, suppose that an airliner cruises at speed v relative to the air. (*a*) For a flight whose one-way distance is d, write an expression for the interval Δt_{calm} needed for a round trip on a windless day. Ignore any time spent on the ground, and assume that the airliner flies at cruising speed for essentially the whole trip. (*b*) Now assume there is a wind of speed w. It doesn't matter which way the wind is blowing; all that matters is that it is a head wind in one direction and a tail wind in the opposite direction. Write an expression for the time interval Δt_{wind} needed for a round trip on the day this wind is blowing. Then show that your expression reduces to

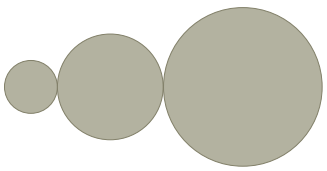
$$\Delta t_{\text{wind}} \approx \Delta t_{\text{calm}}\left[1 + \left(\frac{w}{v}\right)^2\right]$$

provided that $w \ll v$. (For the last step, you may need to know that $(1 \pm z)^b \approx 1 \pm bz$ when $z \ll 1$.) •••

6.6 Center of mass

36. The inertia of an object is m measured when the object is at rest in the Earth reference frame. According to Galilean relativity, what is its inertia measured by an observer moving past the object with a constant velocity $\vec{v}$ in the positive x direction? What is the object's momentum according to this observer? •
37. Jupiter, with an inertia 317.83 times that of Earth, is at an average distance of 7.784×10^{11} m from the Sun. At what distance from the centre of the Sun is the center of mass of the Sun-Jupiter system if the Sun's inertia is 333,000 times that of Earth, and its mean equatorial diameter is $\approx 1.392 \times 10^9$ m? •
38. Figure P6.38 shows three solid spheres of radii a, $2a$, and $3a$ made of materials with densities ρ, 2ρ, and 3ρ. Find the position of the center of mass of the system. ••

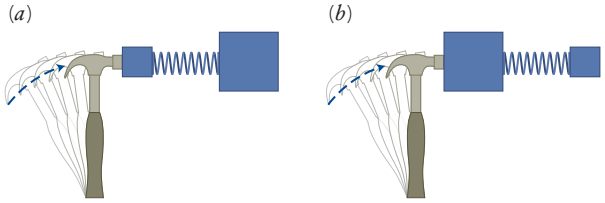
Figure P6.38



39. A boy and a girl are resting on separate rafts 10 m apart in calm water when the girl notices a small beach toy floating midway between the rafts. The girl and her raft have twice the inertia of the boy and his raft. The rafts are connected by a rope 12 m long, so she decides to pull on the rope, drawing the rafts together until she can reach the toy. Which raft gets to the toy first? How much distance is there between the two rafts when the first one reaches the toy? (Ignore the widths of the rafts.) ••

40. The two cubes in Figure P6.40 have different inertias. The cubes are connected to each other by a spring, and a hammer strikes them in the two ways, (*a*) and (*b*), shown in the figure. Assuming that the same impulse is transferred from the hammer, does the center-of-mass motion after the collision depend on which cube is struck first? ••

Figure P6.40



41. Determine the position of the center of mass of the baton shown in Figure P6.41, taking the origin of your coordinate axis to be (*a*) the center of the larger ball, (*b*) the center of the smaller ball, and (*c*) a point 1.0 m to the left of the larger ball. How much calculation was required for each of the three parts of this problem? ••

Figure P6.41

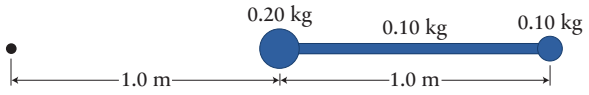
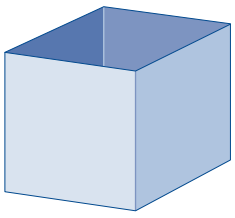


Figure P6.42

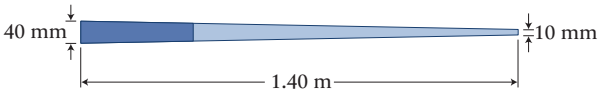


42. The empty cubical box shown in Figure P6.42 has no top face; that is, the box is made up of only five square faces. If all five faces have the same inertia, at what height above the bottom of the box is the center of mass? ••

43. How can you tell from the motion of the center of mass of an isolated system whether the reference frame from which the motion is measured is inertial? ••

44. Determine the center of mass of a pool cue whose diameter decreases smoothly from 40 mm to 10 mm over its 1.40-m length (Figure P6.44). Assume that the cue is made from solid wood, with no hidden weights inside. (Hint: See Appendix D for the center-of-mass computation for extended objects. You will find it easier to do the integral for a complete cone. The pool cue is a truncated cone—that is, a cone with its conical tip removed. Slicing off a piece is like adding negative inertia.) •••

Figure P6.44



6.7 Convertible kinetic energy

45. A bullet of speed v and inertia m strikes and gets embedded in a wooden block of inertia M, initially at rest on a low-friction floor. Plot k_f/k_i as a function of M/m from $M/m = 0$ to $M/m = 10$, and discuss your result. •

46. A 3.0-g particle is moving toward a stationary 7.0-g particle at 3.0 m/s. What percentage of the original kinetic energy is convertible to internal energy? •

47. Think of a system of two objects of different inertias $m_1 < m_2$. How much larger than m_1 can the reduced inertia μ of the pair be? How much smaller than m_1 can it be? •

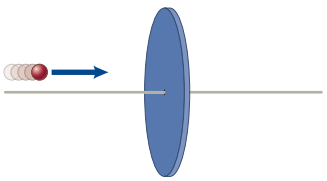
48. (*a*) Is there a reference frame in which the kinetic energy of a system is a minimum? If so, what is this reference frame? (*b*) Is there a reference frame in which the kinetic energy of a system is a maximum? If so, what is this reference frame? ••

49. You hit a pitched baseball with a bat. In which reference frame is the translational (nonconvertible) kinetic energy greater: the reference frame in which the bat is at rest immediately before the collision or the reference frame in which the ball is at rest immediately before the collision? ••

50. What is the ratio K_1/K_2 when K_1 is the kinetic energy converted to internal energy when two cars each initially traveling at 88 km/h collide head-on and K_2 is the kinetic energy converted to internal energy when a car moving at 88 km/h hits a stationary car? Assume that both cars have the same inertia and that the collisions are totally inelastic. (Hint: Determine how much of the energy is required to move the center of mass in each case.) ••

51. A 0.075-kg disk initially at rest in the Earth reference frame is free to move parallel to a horizontal bar through a hole at the disk's center. The disk is struck face-on by a 0.050-kg paintball traveling at 11 m/s, as illustrated in Figure P6.51. (*a*) What is the initial relative velocity of the disk and the paintball? (*b*) What is their reduced inertia? (*c*) If all the paint stays on the disk, what percentage of the original kinetic energy is converted to internal energy? (*d*) What is the change in kinetic energy measured by someone riding a bicycle at the velocity of the paintball before the collision? (*e*) Confirm your answer to part *d* by calculating the final and initial kinetic energies in the cyclist's reference frame. ••

Figure P6.51



52. Take the common case where a moving object of inertia m_{moving} collides with a stationary object of inertia m_{rest}. (*a*) Show that the fraction of kinetic energy not convertible in the collision is $m_{\text{moving}}/(m_{\text{moving}} + m_{\text{rest}})$. Interpret for the case where $m_{\text{rest}} \gg m_{\text{moving}}$ and for the case where $m_{\text{rest}} \ll m_{\text{moving}}$. (*b*) Does the value of the fraction depend on the elasticity of the collision? (*c*) Why is this energy not convertible? ●●

53. Ball 1 is moving toward you at 10 m/s, and you decide to throw ball 2 at it to make it reverse its velocity. The balls collide head-on, and the coefficient of restitution for the collision is 0.90. (*a*) If ball 1 has an inertia of 0.500 kg and ball 2 has an inertia of 0.600 kg, how fast must ball 2 be traveling in order to reverse the velocity of ball 1? (*b*) What is the initial relative velocity of the two balls? (*c*) What is their reduced inertia? (*d*) What percentage of the original kinetic energy is convertible? (*e*) What are the velocities of the two balls immediately after the collision? ●●

54. Ball 1 from Problem 53 is moving away from you at 5.0 m/s, and you decide to throw ball 2 at it to make it go faster. Once again the balls collide head-on, and the coefficient of restitution for the collision is 0.90. (*a*) Given the inertias in Problem 53, how fast must ball 2 be traveling in order to double the velocity of ball 1? (*b*) What is the initial relative velocity of the two balls? (*c*) What is their reduced inertia? (*d*) What percentage of the original kinetic energy is convertible? (*e*) What are the final velocities of the two balls immediately after the collision? ●●

55. If the relative velocities are the same in both cases, which has more convertible kinetic energy: a collision between two objects each of inertia m or a collision between an object of inertia m_1 and an object of inertia $m_2 < m_1$ with $m_1 + m_2 = 2m$? ●●

56. After coming down a slope, a 60-kg skier is coasting northward on a level, snowy surface at a constant 15 m/s. Her 5.0-kg cat, initially running southward at 3.8 m/s, leaps into her arms, and she catches it. (*a*) Determine the amount of kinetic energy converted to internal energy in the Earth reference frame. (*b*) What is the velocity, measured in the Earth reference frame, of an inertial reference frame in which the cat's kinetic energy does not change? ●●

57. You toss a 0.40-kg ball at 9.0 m/s to a 14-kg dog standing on an iced-over pond. The dog catches the ball and begins to slide on the ice. (*a*) Measured from the Earth reference frame, what is the velocity of the dog immediately after he catches the ball? (*b*) Measured from the Earth reference frame, what is the velocity of an inertial reference frame in which the ball's kinetic energy does not change? (*c*) Measured from the Earth reference frame, how much of the original kinetic energy of the system is convertible? (*d*) Measured from the reference frame described in part *b*, how much of the original kinetic energy of the system is convertible? ●●

58. Instead of defining a reduced inertia μ to characterize the convertible kinetic energy of a system, we could define a *reduced velocity* v_{red} as follows: For a system of two particles, one of inertia m_1 and velocity v_1 and the other of inertia m_2 and velocity v_2, the kinetic energy is

$$K = K_{\text{cm}} + K_{\text{conv}} = \tfrac{1}{2}(m_1 + m_2)v_{\text{cm}}^2 + \tfrac{1}{2}(m_1 + m_2)v_{\text{red}}^2.$$

Derive an expression for the reduced velocity in terms of the given variables. (Just as the center-of-mass velocity is a weighted mean of the individual velocities, the reduced velocity is a weighted difference between the individual velocities.) ●●●

6.8 Conservation laws and relativity

59. In this chapter you learned about the zero-momentum reference frame. You may be curious whether there is such a thing as a zero-energy reference frame. Does a zero-kinetic energy reference frame always exist, never exist, or sometimes exist? ●

60. After a totally inelastic collision, the kinetic energy of an isolated system composed of two objects is zero. What was the momentum of the system in the same reference frame before the collision? ●

61. A mother penguin and her chick are on a flat, icy surface. The mother is lying at rest 0.50 m from the edge of the water. The chick, which has one-fourth of its mother's inertia, is sliding, collides with her inelastically, and bounces back at one-eighth of its original speed. The mother wakes up as she hits the water 0.40 s later. (*a*) How fast was the chick going when it hit her? (*b*) To a penguin paddling directly toward them at 1.0 m/s, what is the mother's momentum before and after the collision in terms of m_{mother}? ●●

62. A 50-kg ice skater moves across the ice at a constant speed of 2.0 m/s. She is caught by her 70-kg partner, and then the pair continues to glide together. He is at rest when he catches her, and immediately afterward they both coast. (*a*) What is their velocity just after the catch? (*b*) A skater follows the girl at a constant speed of 1.0 m/s. What is the girl's change in momentum in the reference frame of the skater? ●●

63. In an inertial reference frame F, an orange of inertia m_{orange} and velocity $\vec{v}_{\text{orange}}$ collides totally inelastically with an apple of inertia m_{apple} initially at rest. (*a*) When this collision is viewed from an inertial reference frame G, the kinetic energy of the orange is unchanged by the collision. What does an observer in the F reference frame measure for the velocity of the G reference frame? (*b*) For the case where $m_{\text{orange}} = m_{\text{apple}}$, sketch the velocity vectors of the two objects in each reference frame before and after the collision. ●●

64. An 800-kg car moving at 108 km/hr hits and gets entangled with a large truck of inertia 10,000 kg initially at rest. What is the kinetic energy of the wreckage? How much energy is lost from the system? ●●

65. A 0.20-kg softball is traveling at a velocity of 20 m/s to the east relative to Earth. It collides head-on with a 0.40-kg rubber ball traveling at a velocity of 10 m/s to the west. (*a*) If the system's kinetic energy, as measured from the Earth reference frame, decreases by 20% because of the collision, what are the final velocities of the balls? (*b*) What change in internal energy has occurred? (*c*) An observer watches this collision from a reference frame moving at a velocity of 15 m/s to the east relative to the Earth reference frame. What changes in kinetic and internal energies does this observer measure? (*d*) What

changes in kinetic and internal energies would be measured by an observer in a reference frame traveling at 20 m/s east relative to the Earth reference frame? ●●

66. Derive an expression showing that when an elastic collision between two objects is viewed from the zero-momentum reference frame, the direction of the momentum of each object is reversed and the magnitude of the change in momentum of each object is twice the magnitude of that object's initial momentum. ●●●

Additional Problems

67. Just as a car passes a school crossing guard, a child throws a toy from the back seat of the car toward his sister in the front seat. The toy is thrown at a speed of 2.0 m/s relative to the car, and the car is traveling at 10 m/s relative to Earth. Draw the velocity vectors for the toy, the car, and the crossing guard from the reference frame of (*a*) the car, (*b*) the guard, and (*c*) the toy. ●

68. A 15-kg box is pushed at 5 m/s across a low-friction floor and is caught by a 60-kg man on low-friction skates, initially at rest. What is the resulting kinetic energy of the man? How much energy is convertible? ●

69. A 72-kg woman is walking at 1.5 m/s. An 8-kg dog is running at six times that speed in the same direction. At what speed and in what direction relative to the woman would you have to be jogging in order for the dog to have the same momentum as the woman in your reference frame? ●

70. One rider is in a descending elevator that is accelerating to a stop. Another is in an elevator that is accelerating upward from rest. With their eyes closed, can the riders tell which of the two elevators they are in? ●●

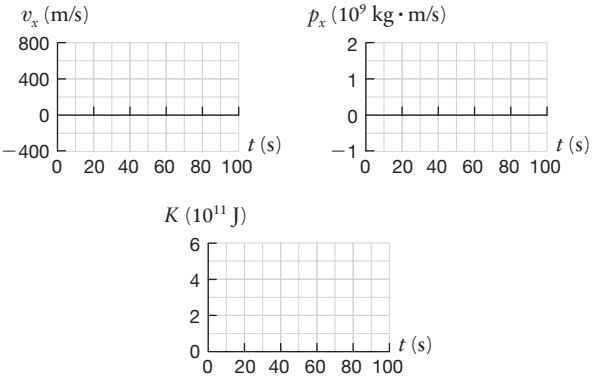
71. You are in a noninertial reference frame observing two isolated objects of inertias m_1 and m_2 that move at constant speed relative to each other. What can you say about each object's (*a*) apparent acceleration and (*b*) apparent change in momentum? ●●

72. A bug of inertia m_B collides with the windshield of a Mack truck of inertia $m_T \gg m_B$ at an instant when the relative speed of the two is v_{BT}. (*a*) Express the system momentum in the truck's reference frame, then transform that expression to the bug's reference frame, and in so doing remove $m_B v_{BT}$ from the expression. (Remember, in the bug's reference frame, the bug is initially at rest and the truck is moving.) (*b*) Now express the system momentum in the bug's reference frame, then transform that expression to the truck's reference frame, and in so doing remove $m_T v_{BT}$ from the expression. (*c*) Is there something wrong here? How can we change the momentum by a small amount $m_B v_{BT}$ doing the transformation one way and by a large amount $m_T v_{BT}$ doing the transformation the other way? ●●

73. A 0.045-kg golf ball moving at 50 m/s (measured in the Earth reference frame) collides inelastically with a 1.8-kg, heavy-duty plastic flowerpot sitting on a windowsill. The coefficient of restitution for the collision is 0.50. Calculate the final velocities of the ball and the pot. Do this problem by first transforming to the zero-momentum reference frame, where the collision is much simpler, and then transforming back to the Earth reference frame. (Assume that all the action happens in one dimension and that the friction between pot and sill is insignificant.) ●●

74. Asteroid A1, $m_{A1} = 3.60 \times 10^6$ kg, and asteroid A2, $m_{A2} = 1.20 \times 10^6$ kg, collide head-on in space. Approximate (rather poorly) the collision as being elastic. Observers watch the event from two space platforms. An observer on platform Q measures the speed of A1 to be 528 m/s and the speed of A2 to be 315 m/s, and sees A2 heading directly toward A1. The second platform, platform Z, is at rest in the zero-momentum reference frame of the asteroids. (*a*) What is the combined momentum of the asteroids measured from platform Q? (*b*) What is the velocity of platform Z relative to platform Q? (*c*) Draw velocity, momentum, and kinetic energy graphs for the asteroids as viewed from platform Z, using graphs like those in Figure P6.74. Assume the collision begins at $t = 40$ s and lasts for 20 s. (*d*) Add a curve for the momentum of the system to the momentum graph you constructed in part *c*. (*e*) What happens to the kinetic energy when its curve bends? (*f*) On your kinetic energy graph from part *c*, show what an observer on platform Z sees for the convertible and nonconvertible kinetic energies of the asteroids. (*g*) Repeat parts *c*–*f* as seen from platform Q. ●●

Figure P6.74



75. Assume that the asteroids in Problem 74 collide totally inelastically. Answer the same questions posed in that problem for this situation. ●●

76. A 1500-kg van is coasting to a stoplight at 15 m/s. A 1000-kg car behind the van is doing the same thing at 25 m/s and crashes into the rear of the van. The bumpers collide with a coefficient of restitution of 0.70. (*a*) What are the final velocities of the two vehicles? (*b*) How much kinetic energy is converted to internal energy in this collision? (*c*) How much energy is converted if the bumpers lock together in a totally inelastic collision? ●●

77. A 0.30-kg cart traveling along a low-friction track at 2.0 m/s relative to Earth collides with a 0.50-kg cart traveling in the same direction at 1.0 m/s relative to Earth. If the system's kinetic energy measured in the Earth reference frame increases by 30% as a result of the explosive collision, what are the final velocities of the carts in this reference frame? ●●

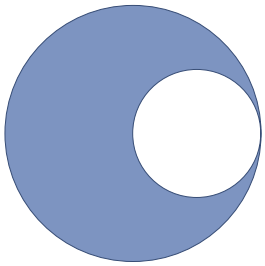
78. The transformation between position and time measurements in an inertial reference frame I and position and time measurements in a constantly accelerating (noninertial) reference frame N is given by

$$t_N = t_I = t$$
$$\vec{r}_N = \vec{r}_I - \vec{v}_{IN}t - \tfrac{1}{2}\vec{a}_{IN}t^2,$$

where $\vec{v}_{IN}$ is the velocity of the noninertial reference frame at $t = 0$ measured from the inertial reference frame and $\vec{a}_{IN}$ is the acceleration of the noninertial reference frame measured from the inertial reference frame. Derive by differentiation the transformation laws, giving v_N and a_N in terms of v_I and a_I. ●●●

79. The medallion shown in Figure P6.79 has been made by cutting a small circular piece out of a larger circular disk. The diameter of the original disk is twice the diameter of the hole, and the thickness of the disk is uniform. You begin to wonder about the location of the center of mass of this object, but the calculation seems impossibly difficult. Then you visualize two circular pieces coming together in superposition. Where is the center of mass? ●●● CR

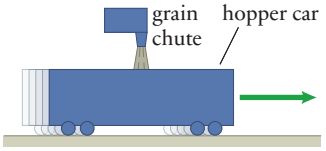
Figure P6.79



80. As an avid biker, you've come to realize that your top speed is really an air speed, not a ground speed, because air resistance is very noticeable when you ride fast. You like to ride at top speed for exactly an hour every day, and you know that on a calm day, you can ride 20 km before having to turn around and come back the same way. Today, though, you have a 20-km/h head wind for the first leg of the trip (which of course is a tailwind on the way back). You must be back home in an hour, so you look at your map. What top speed can you achieve today? ●●● CR

81. An uncoupled 20,000-kg railroad hopper car coasts along a track at 2.0 m/s. As it passes a grain chute, the chute opens (Figure P6.81), grain fills the car at a rate of 4000 kg/s for 5.0 s, and the chute closes. (*a*) How fast is the car going after the chute closes? (*b*) How far does the car travel during the 5.0 s in which grain is being added? (This distance is the length of the trail of grain in the car.) (*c*) Where is the center of mass of the grain-car system at the end of the 5.0 s? (*d*) Where is the center of mass of the trail of grain at the end of the 5.0 s? (*e*) Why doesn't the center of mass of the trail of grain coincide with the point midway between the two ends of the trail? ●●●

Figure P6.81



82. A skateboarder coasting at 5.0 m/s decides to coast into a friend going 4.0 m/s on rollerblades in the same direction. Unfortunately, the rollerblader stops right before the collision, too quickly for the skateboarder to react. The rollerblader turns around as he stops and pushes his arms out to ward off the collision. The rollerblader and his gear have a combined inertia of 70 kg, and the skateboarder and his gear have a combined inertia of 79 kg. (*a*) If the skateboarder ends up moving backward at 2.0 m/s, what is the coefficient of restitution for the collision? (*b*) What are the changes in kinetic and internal energies measured by an observer standing alongside the collision? (*c*) What are the changes in kinetic and internal energies measured by an observer in the zero-momentum reference frame for the skateboarder-rollerblader system? (*d*) If the rollerblader had not turned around and stopped, so that the skateboarder was able to grab him and keep coasting, what amount of kinetic energy would have been lost in the collision? ●●●



7

Interactions

7.1 The effects of interactions
7.2 Potential energy
7.3 Energy dissipation
7.4 Source energy
7.5 Interaction range
7.6 Fundamental interactions

7.7 Interactions and accelerations
7.8 Nondissipative interactions
7.9 Potential energy near Earth's surface
7.10 Dissipative interactions

CONCEPTS

QUANTITATIVE TOOLS

In past chapters we saw how interactions can change the kinetic energy of objects as well as their internal energy. Every event that happens in this universe is the result of some interaction between objects. Interactions determine the structure of the universe, from the subatomic scale to the cosmological scale. In this chapter we study how interactions convert energy from one form to another. In the process we learn more about the concept of internal energy introduced in Chapter 5.

7.1 The effects of interactions

In the broadest sense, **interactions** are mutual influences between two objects that produce change (either physical change or a change in motion). As an example, consider a cart at rest on a horizontal track. The only way to get the cart to move is to make it *interact* with something else. We could, for instance, give it a shove or make it collide with another cart or fasten a magnet to it and push or pull it with another magnet.

Note that in each case we make the cart interact with something else. We can't make it change its state of motion without that something else: a hand, another cart, a magnet. If you look at Figure 4.19, you'll see that in each case where there is an interaction (*a*, *c*, and *d*), the interaction takes place between *two* objects: puck and floor, cart and cart, hand and cart.

An interaction that causes objects to accelerate can be either repulsive or attractive. A *repulsive interaction* is one in which the interacting bodies accelerate away from each other; an *attractive interaction* is one in which the interacting bodies accelerate toward each other. Some interactions are repulsive under some conditions and attractive under others, as illustrated in **Figure 7.1**.

CONCEPTS

7.1 (*a*) Imagine holding a ball a certain height above the ground. If you let the ball go, it accelerates downward. An interaction between the ball and what other object causes this acceleration? Is this interaction attractive or repulsive? (*b*) Once the ball hits the ground, its direction of travel reverses. Is this reversal the result of an attractive interaction or a repulsive one?

Figure 7.1 An interaction between two carts linked by a spring. When the spring is stretched, the interaction is attractive; when the spring is compressed, the interaction is repulsive.

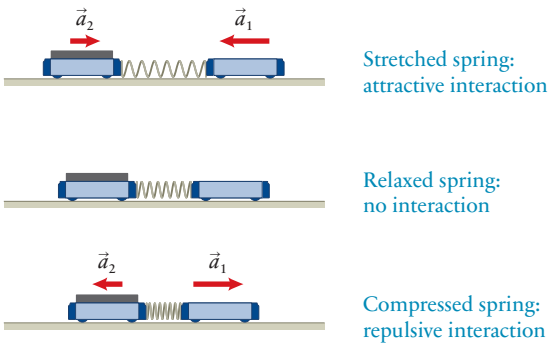


Figure 7.2 shows the x components of the velocities, momenta, and accelerations as well as the kinetic energies of two carts before and after an interaction. Cart 1, of inertia $m_1 = 0.12$ kg and initially at rest, is hit by cart 2, of inertia $m_2 = 0.24$ kg and initially moving at 0.55 m/s. The collision is elastic.

Figure 7.2*a* shows the x components of the velocities of the two carts. We've seen similar graphs before (Figure 4.8, for example). The x components of the final velocities are +0.73 m/s for cart 1 and +0.18 m/s for cart 2, and their relative speed remains 0.55 m/s, as it must for an elastic collision.

Figure 7.2*b* shows what happens to the x components of the momenta: One cart undergoes a momentum change in one direction, and the other cart undergoes an equal momentum change in the opposite direction. This is true at every instant: Even *during* the interaction, the momentum of the two-cart system is constant.

7.2 (*a*) Use Figure 7.2*a* to calculate the x components of the momenta of the two carts at $t = 30$, 60, and 90 ms. (*b*) What is the x component of the momentum of the system at each of these three instants?

Notice in Figure 7.2 that the interaction that causes the changes in velocity and momentum is not instantaneous; instead, it takes place over a certain time interval (represented by the shaded area in the graphs). The x component of the average acceleration of each cart during the interaction is equal to the change in the x component of the velocity for that cart divided by the duration of the interaction. Because the change in speed for cart 1 is twice as big as that for cart 2 (Figure 7.2*a*) and because the time interval over which the interaction takes place is the same for both carts, the magnitude of the average acceleration of cart 1 is twice that of cart 2. This result is a direct consequence of conservation of momentum.

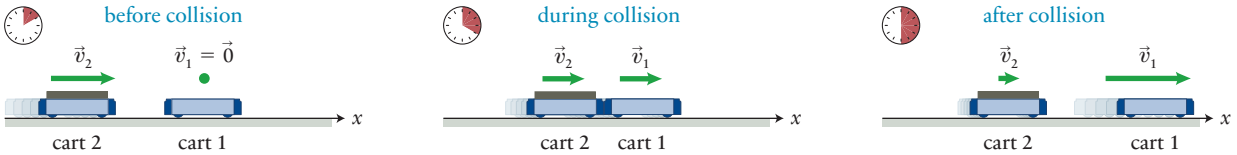
As we shall see in Section 7.7, this relationship between the average accelerations of colliding bodies is true for the instantaneous accelerations as well. So, whenever two objects interact, the ratio of the x components of their accelerations is equal to the negative inverse of the ratio of their inertias.

EXAMPLE 7.1 Crash

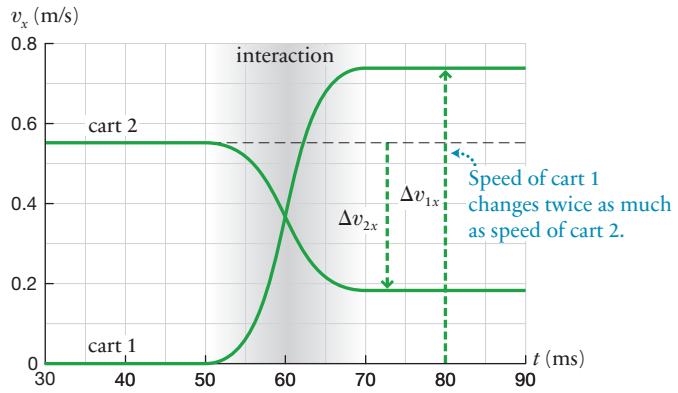
A small car and a heavy truck moving at equal speeds in opposite directions collide head-on in a totally inelastic collision. Compare the magnitudes of (*a*) the changes in momentum and (*b*) the average accelerations of the car and the truck.

1 GETTING STARTED I begin by making a sketch of the situation before and after the collision (**Figure 7.3**). Before the collision, the truck and car both move at the same speed. After the totally inelastic collision, the two move as one unit with zero relative velocity. Because the inertia m_t of the truck is greater than the inertia m_c of the car, the momentum of the system points in the same direction as the direction of travel of the truck. The combined wreck must therefore move in the same direction after the collision.

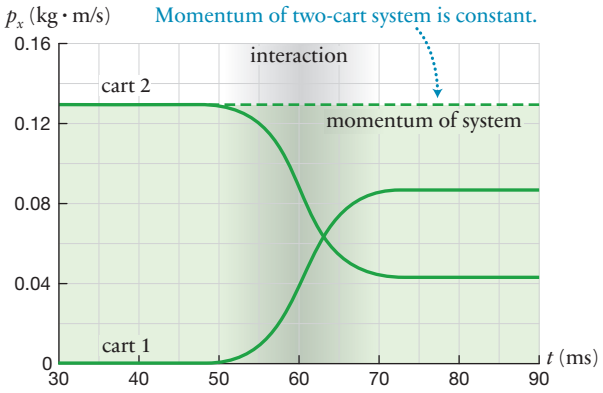
Figure 7.2 Conservation of momentum and kinetic energy in an elastic collision between two carts on a low-friction track. The inertia of cart 1 is 0.12 kg, and that of cart 2 is 0.24 kg.



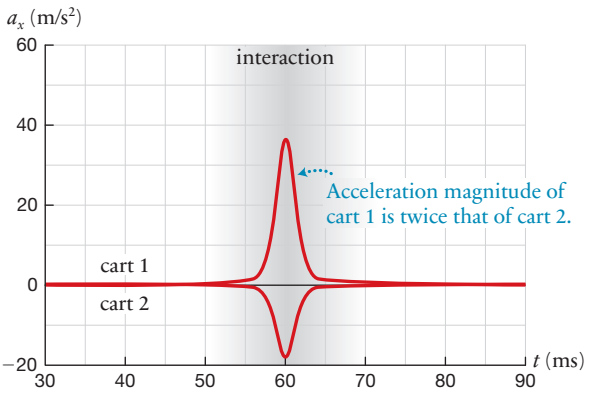
(*a*) Velocity



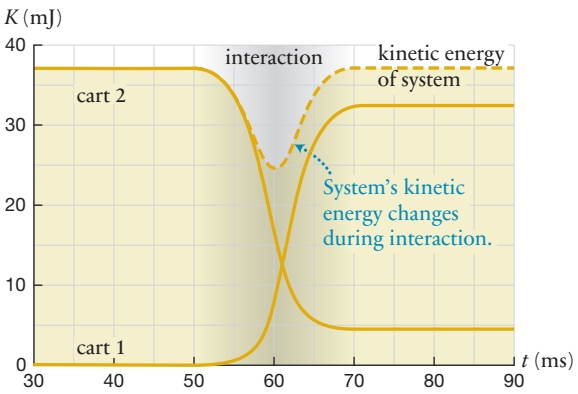
(*b*) Momentum



(*c*) Acceleration

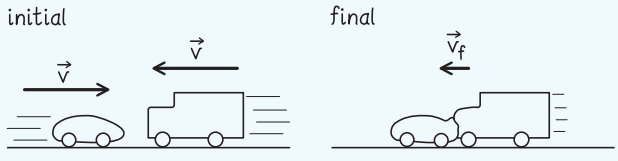


(*d*) Kinetic energy



2 DEVISE PLAN To compare the changes in momentum, I can apply conservation of momentum to the isolated truck-car system. I can obtain the change in velocity by dividing the change in momentum by the inertia. Because the changes in velocity occur over the same time interval for both car and truck, and because the magnitude of the average acceleration is given by $\Delta v/\Delta t$, the ratio of the accelerations is the same as the ratio of the changes in velocity.

Figure 7.3 Example 7.1.



3 EXECUTE PLAN (*a*) The momentum of the isolated truck-car system does not change in the collision, and so the magnitudes of the changes in momentum for the car and the truck are the same. ✔

(*b*) The change in the x component of the velocity of the truck is $\Delta p_{tx}/m_t$, and the change in the x component of the velocity of the car is $\Delta p_{cx}/m_c$. Because $m_t > m_c$ and because the magnitudes of the changes in momentum are equal, I conclude that the magnitude of the velocity change of the car is larger than that of the truck. ✔

4 EVALUATE RESULT In any collision the magnitudes of the changes in momentum are the same for the two colliding objects, and so the answer to part *a* does not surprise me. That the magnitude of the velocity change for the car is larger also makes sense: As my sketch shows, the velocity of the car reverses, whereas the truck slows down somewhat but keeps traveling in the same direction.

CONCEPTS

Figure 7.2*d* shows the kinetic energy of the two carts. Just like the momentum, the kinetic energy of each cart changes, but because the collision is elastic, the kinetic energy of the system before the interaction is the same as the kinetic energy of the system afterward. Note, however, that the kinetic energy of the system does not remain constant *during* the collision! Unlike momentum, the kinetic energy of a system of colliding objects changes during the interaction—even when the interaction is an elastic collision.

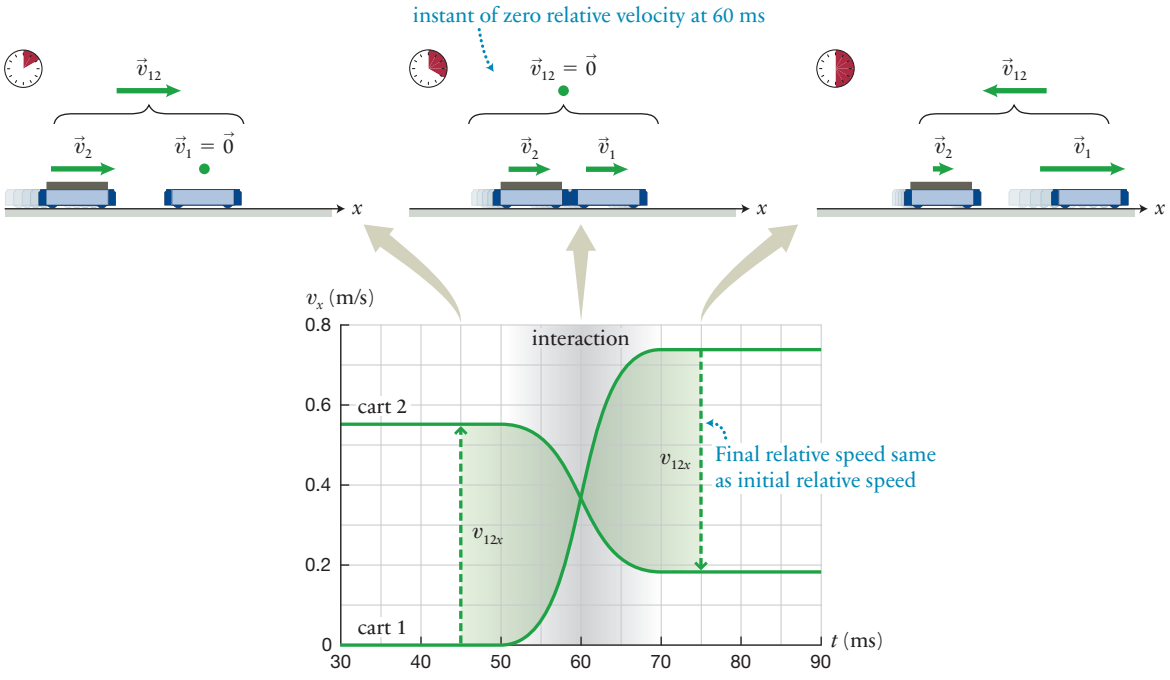
To see why, remember from Section 5.5 that kinetic energy's being constant is mathematically equivalent to saying that the relative speed of two objects is the same before and after an interaction. Whenever two objects collide, however, their relative speed has to change during the interaction. If the objects were to maintain a constant relative speed during the interaction, changing their velocities would change the momentum of the system, and this can't happen if the system is isolated. Consequently the kinetic energy of any system of colliding objects has to change during the interaction, regardless of the type of collision—elastic, inelastic, or totally inelastic.

Figure 7.4 illustrates the instant of zero relative velocity for the elastic collision shown in Figure 7.2. At the instant at which the two $v(t)$ curves intersect ($t = 60$ ms), the velocities are equal, and hence the relative velocity is zero at that instant. If you cover up the rightmost diagram at the top of Figure 7.2 (representing the carts' motion after the collision), you have no way of distinguishing this elastic collision from a *totally inelastic* one. In other words, in an elastic collision and a totally inelastic collision that have the same initial conditions, the same amount of kinetic energy is converted. The difference is that in the elastic collision any kinetic energy in the system before the collision that is converted to internal energy during the collision is all converted back to kinetic energy after the collision, whereas in the totally inelastic collision all of the initial kinetic energy converted to internal energy during the collision stays as internal energy and is not converted back to kinetic energy. The important point here is that, whenever two objects interact, their relative speed has to change, and therefore the kinetic energy of the system must change during the interaction.

7.3 (*a*) Use Figure 7.2*a* to calculate the kinetic energies of the two carts at $t = 30$, 60, and 90 ms. (*b*) What is the kinetic energy of the system at each instant?

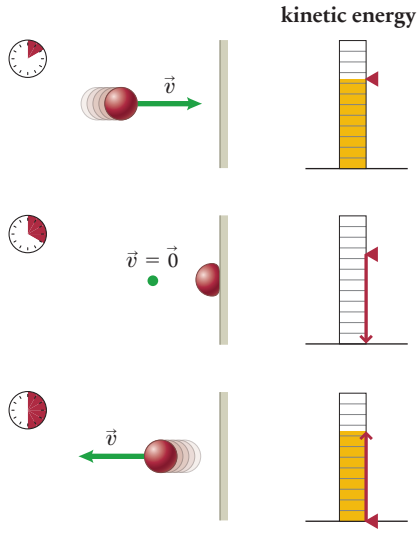
This disappearance of kinetic energy doesn't jeopardize conservation of energy. The kinetic energy "missing" during the interaction has merely been temporarily converted to internal energy. Consider, for instance, a rubber ball thrown against a wall and bouncing off, as in **Figure 7.5**. At the instant of impact, the ball has zero velocity and has lost all the kinetic energy it had while moving toward the wall. Where has all this kinetic energy gone? It has gone into changing the shape of the ball. **Figure 7.6** shows a golf club striking a golf ball. The deformation of the ball at the instant of impact is

Figure 7.4 Even in an elastic collision between two carts, there is an instant at which the carts have the same velocity, which means that at that instant their relative velocity is zero.



CONCEPTS

Figure 7.5 At the instant the ball in this elastic collision has zero velocity, all its kinetic energy has been used to deform the ball (and, to a lesser extent, the wall). Thus, this energy is stored temporarily as elastic potential energy.



clearly visible. After impact, both the rubber ball of Figure 7.5 and the golf ball of Figure 7.6 regain their original shape, and the kinetic energy converted to internal energy during the interaction reappears as kinetic energy. Because the system (ball and wall in Figure 7.5, ball and club in Figure 7.6) is closed, the energy of the system remains constant throughout the collision. In the elastic collision in Figure 7.4, similar changes in shape take place in the carts, but these changes occur too fast and are too small to be observed easily.

In inelastic collisions, the system's kinetic energy is not the same before and after the collision. As in an elastic collision, there is an instant at which the relative speed of the colliding objects is zero, but unlike in an elastic collision, not all of the kinetic energy converted to internal energy during the interaction reappears as kinetic energy after the collision. Some of the kinetic energy is converted irreversibly to internal energy (see Section 5.3).

We can summarize the characteristics of an interaction:

1. *Two objects* are needed.
2. The *momentum* of a system of interacting objects is the same before, during, and after the interaction (provided the system is isolated).

Moreover, for interactions that affect the motion of objects:

1. The ratio of the x components of the accelerations of the interacting objects is the negative inverse ratio of their inertias. Because the velocities of both objects change in an interaction, the *individual* momenta and kinetic energies change.
2. The system's *kinetic energy changes during the interaction*. Part of it is converted to (or from) some internal energy. In an elastic collision, all of the converted energy reappears as kinetic energy *after* the collision. In an inelastic collision, some of the converted kinetic energy reappears as kinetic energy; in a totally inelastic collision, none of the converted kinetic energy reappears.

7.4 (*a*) In Figure 7.5, what is the momentum of the ball during the collision? (*b*) Is the momentum of the ball constant before, during, and after the collision? If so, why? If not, why not, and for what system is the momentum constant?

Figure 7.6 At the instant of impact, a golf club and golf ball have zero relative speed, and much of the system's initial kinetic energy has gone into severe deformation of the ball (plus some deformation of the elastic club).

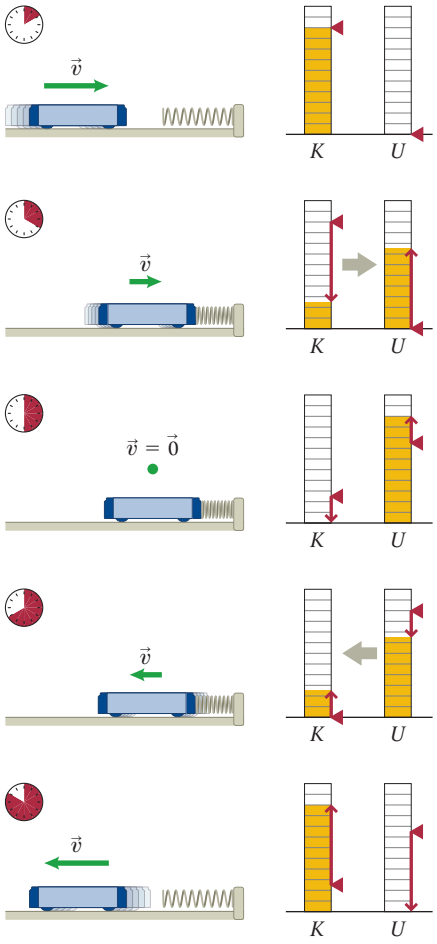


7.2 Potential energy

In any collision or interaction, the part of the converted kinetic energy that is temporarily stored in reversible changes in physical state and is then converted back to kinetic energy after the interaction is called **potential energy,** represented by U. Potential energy is a form of internal energy. The term *potential* refers to the fact that the energy has the potential to be converted back to kinetic energy—it is only temporarily stored. Just think of the rubber ball in Figure 7.5. At the instant it has zero velocity, all the energy that before impact was kinetic energy has gone into deforming the ball and is therefore now stored inside the ball. Because rubber is elastic, the stored energy has the potential to revert to kinetic energy. Similarly, if a cart on a track hits a spring mounted on a post (**Figure 7.7** on the next page), the cart compresses the spring, and all the energy that before impact was kinetic energy of the cart is temporarily converted to potential energy stored in the spring. As the spring returns to its original shape, this potential energy is converted back to kinetic energy of the cart. Throughout the interaction between the cart

CONCEPTS

Figure 7.7 Energy conversion during the compression of a spring by a moving cart.



and the spring, the sum of the kinetic and potential energies of the closed cart-spring system is constant.

How and where is energy stored as potential energy? It is stored in reversible changes in the *configuration state* of the system, by which I mean the spatial arrangement of the system's interacting components.

There are many forms of potential energy, all related to the way interacting objects arrange themselves spatially. If you hoist a boulder up through a third-floor window, you change the configuration state of the boulder-Earth system. As the boulder is hoisted upward, the form of potential energy called **gravitational potential energy** is stored in the Earth-boulder system. To see that the stored energy is potential energy, imagine cutting the rope that holds the boulder in place: The boulder accelerates as it tumbles back to Earth because of the gravitational interaction between it and Earth. As the boulder gains speed, its kinetic energy increases. Because energy is conserved and because the boulder-Earth system is closed, the gain in kinetic energy has to be compensated by a loss of potential energy associated with the changed configuration state.

Moving to the atomic level, whenever you squeeze a ball or a spring, you change the configuration state of the atoms that make up the ball or spring. Reversible deformations correspond to changes in **elastic potential energy.** In Figure 7.7, for instance, the kinetic energy of the cart is temporarily converted to elastic potential energy in the spring.

> **Potential energy is the form of internal energy associated with reversible changes in the configuration state of an object or system. Potential energy can be converted entirely to kinetic energy.**

EXERCISE 7.2 Launch

A ball is pressed down on a spring and then released from rest. The spring launches the ball upward. Identify the energy conversions that occur between the instant the ball is released and the instant it reaches the highest point of its trajectory.

SOLUTION As the spring expands, elastic potential energy stored in the spring is converted to kinetic energy of the ball. As the ball travels upward, it slows down and the kinetic energy is converted to gravitational potential energy of the ball-Earth system. ✔

7.5 In Figure 7.7, consider the cart's initial speed to be v_i. Assuming no potential energy is initially stored in the spring, how much potential energy is stored in the spring at the instant depicted in the middle drawing and at the instant depicted in the bottom drawing? Give your answers in terms of m, v_i, and v.

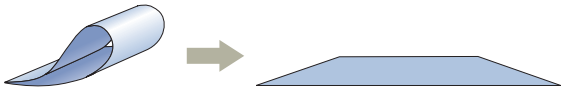
7.3 Energy dissipation

The part of the converted kinetic energy that does not reappear after an inelastic collision is said to be *dissipated* (in other words, irreversibly converted). To understand what happens to energy that is dissipated, try the following simple experiment with a crisp sheet of paper lying flat on a horizontal surface. Without making any folds, bend one end of the sheet over the other, as illustrated in **Figure 7.8*a***. If you remove your hand, the sheet straightens itself out. If you crumple up the sheet, however, as in Figure 7.8*b*, and then remove your hand, the crumpled sheet may expand a little, but it doesn't regain its original shape. There is an important difference between the two changes in shape. In Figure 7.8*a* the deformation takes place in a *coherent* manner, which means that at the atomic level, there is a pattern in the displacements of the atoms: They move orderly in rows, with each successive row experiencing a small displacement in the same direction as adjacent rows. As you bend the sheet in this fashion, you store potential energy in it. When you release the sheet, this potential energy appears as kinetic energy. In Figure 7.8*b*, the change in shape is *incoherent* because the atoms are displaced in random directions.

CONCEPTS

Figure 7.8 Reversible and irreversible deformations of a sheet of paper. If the sheet is simply bent (*a*), it returns spontaneously to its original shape; if it is crumpled (*b*), it does not.

(*a*) Coherent deformation: reversible



(*b*) Incoherent deformation: irreversible



The sheet can't straighten itself out because the randomly displaced atoms are in one another's way. Some of the smaller, less random deformations can restore themselves when you let go, but in general the energy used to do the crumpling has been irreversibly converted, producing a permanent deformation.

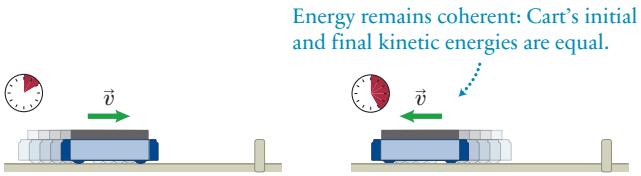
In addition to being dissipated to incoherent deformations, energy can be dissipated to incoherent motion, as shown in **Figure 7.9**. In Figure 7.9*a* a cart with a metal bar on top collides with a post that is anchored to the track. The collision is elastic, and so the cart's final kinetic energy is equal to the initial kinetic energy. In Figure 7.9*b* the metal bar has been replaced by a frame from which a large number of little balls are suspended. When this cart hits the post, the cart comes to a dead stop but the balls—initially at rest relative to the cart—chaotically jiggle back and forth in random directions. What has happened is that the initial kinetic energy of the cart has been converted to kinetic energy associated with the incoherent motion of the balls. We still have exactly the same amount of kinetic energy, but it has all become "internal." The initially coherent kinetic energy of the moving cart has been dissipated to the incoherent kinetic energy of the balls.

The experiment illustrated in Figure 7.9*b* is a large-scale model of what happens at the atomic scale in inelastic collisions. When a tennis ball bounces off a hard surface, part of the ball's kinetic energy is converted to incoherent energy of the atoms that make up the ball. (Imagine the atoms in the ball jiggling incoherently like the balls in Figure 7.9*b*.) Consequently the ball's kinetic energy after the bounce is less than that before the bounce, and so the rebound height is below the original height.

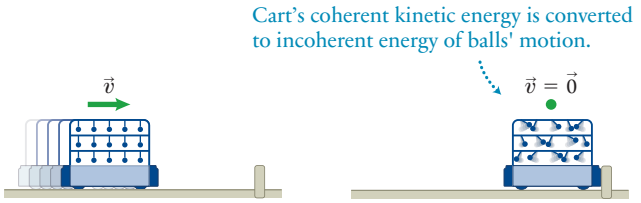
We can now give a complete classification of energy (**Figure 7.10** on the next page). All energy can be divided into two fundamental classes: energy associated with motion and energy associated with the configuration of interacting objects. Each class of energy comes in two forms: coherent and incoherent.

Figure 7.9 Reversible and irreversible collisions. In the totally inelastic collision in (*b*), the cart's coherent kinetic energy is converted completely to random energy of motion in the suspended balls.

(*a*) Cart rebounds elastically from barrier



(*b*) Barrier stops cart, putting balls in random motion



When all the atoms in an object move in a coherent fashion in the same direction—which means the entire object moves in that direction—the energy of motion is called *kinetic energy*. Energy stored in coherent changes in configuration is called *potential energy*. The sum of a system's kinetic energy and potential energy is called the system's **mechanical energy** or **coherent energy.**

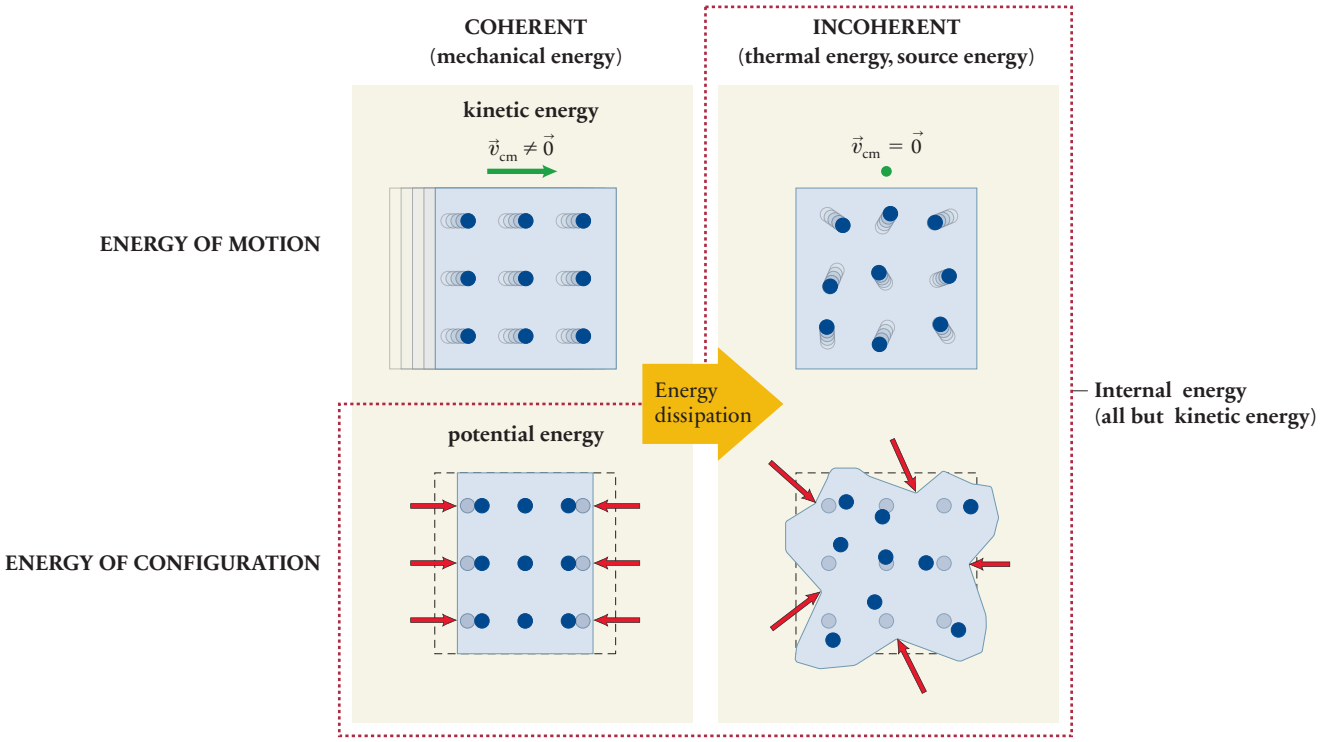
In addition to this coherent energy, a system can have **incoherent energy** associated with the incoherent motion and configuration of its parts. An important part of a system's incoherent energy is its **thermal energy.** (As we shall see in Checkpoint 7.7, not all incoherent energy is thermal energy.) In most cases, the higher the thermal energy of an object, the higher its temperature—a concept we shall return to in Chapter 20. A rubber ball that has been squeezed repeatedly becomes noticeably hotter because the incoherent energy of the atoms that make up the ball has increased.

Here is a quick experiment you should try. Unfold a metal paper clip so that it takes on an S-shape. Holding the two ends, touch your lips with the central part of the clip—it should feel cold. Then vigorously bend the clip back and forth until it snaps in the middle and hold the two broken ends against your lips—the ends feel *hot*. The coherent energy you put into the back-and-forth bending has been converted to incoherent energy of the atoms in the metal (and some has gone into the breaking of the metal). When you hold the broken ends against your lips, you "feel" the atoms in the metal jiggling back and forth—they bounce against the atoms in your lips, transferring some of their energy to the atoms in your lips, thereby warming up your lips.

The sum of the system's incoherent energy and its potential energy is the system's *internal energy*.

CONCEPTS

Figure 7.10 Classification of energy.



Energy dissipation corresponds to the conversion of coherent energy into incoherent energy. This process is essentially irreversible—it is highly improbable that incoherent energy will revert by itself to a coherent form. Consider the cart in Figure 7.9*b*. If you shake it so that all the balls jiggle back and forth wildly, you are storing a certain amount of incoherent energy in the cart. If you then place the cart with the jiggling balls on the track against the post, this incoherent energy doesn't spontaneously turn into coherent kinetic energy of the cart as a whole—the cart does not suddenly start moving away from the post. What does happen, if you wait, is that the motion of the balls eventually stops. The kinetic energy of each ball gets dissipated, too—broken down into the incoherent jiggling of the individual atoms that make up each ball. (Because of the increased jiggling, the temperature of the balls increases but, as we shall see later, the temperature rise is imperceptibly small.)

This breaking down of the coherent energy of a large object into smaller units of incoherent thermal energy raises an interesting question: What happens to the energy once it has become thermal energy? Except for the motion of the Sun, the Moon, and other celestial objects, *all* motion we are familiar with comes to a stop unless something keeps it going. The stopping occurs because dissipation causes the coherent energy of the moving object to turn into incoherent energy. Does the energy of motion of the object's atoms dissipate, too? The answer to this question is *no*; at the atomic level, there simply is no way to further break down and randomize the energy.

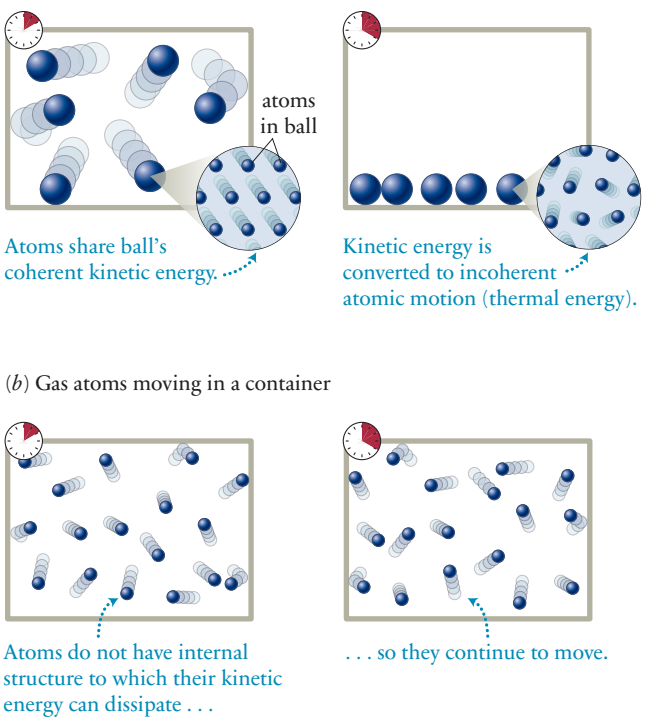
To see that this is so, suppose we throw a bunch of rubber balls into a box (**Figure 7.11*a***). The balls initially bounce around, but if we look into the box a minute or so later, all of them are lying motionless on the bottom. All the energy is still there, but now it is inside the individual balls in the form of thermal energy. (The balls don't feel any warmer to you because the increase in thermal energy is so small as to be imperceptible.) Now suppose we observe a number of gas atoms randomly flying around in a container (Figure 7.11*b*). Provided the atoms do not interact with their environment, they keep flying around in the container *forever* because there are no smaller parts inside the atoms to which their kinetic energy can be dissipated.* And, because there is no dissipation at the atomic scale, nothing is required to keep atoms moving.

Friction, which occurs whenever two solid objects rub against each other or when an object moves through any fluid (air or water, say), is one of the causes of energy dissipation. Friction also causes wear and tear by removing microscopic pieces of material.

*What about the particles that make up the atoms—electrons, protons, neutrons? As we shall see in Chapter 20, these parts can store energy only in certain minimum increments that are larger than the typical kinetic energy of an atom, and so the kinetic energy of an atom cannot be further dissipated inside the atom.

CONCEPTS

Figure 7.11 (*a*) Bouncing balls come to rest because their kinetic energy is converted to internal kinetic energy of the atoms (thermal energy). (*b*) Atoms have no internal structure to which kinetic energy can be dissipated, so they keep moving.



This removal of material involves breaking the chemical bonds that hold atoms together. The incoherent configuration energy associated with the breaking of chemical bonds by friction or by crumpling, however, is many orders of magnitude smaller than the thermal energy generated in these processes (see the box "Coherent versus incoherent energy" on page 221). In general, we assume that the only type of dissipation caused by friction and crumpling is dissipation of mechanical energy into thermal energy. In Chapter 19 we shall examine energy dissipation and the accompanying conversion of coherent energy to incoherent energy in more detail.

7.6 Because of friction, a 0.10-kg hockey puck initially sliding over ice at 8.0 m/s slows down at a constant rate of 1.0 m/s^2 until it comes to a halt. (*a*) On separate graphs, sketch the puck's speed and its kinetic energy as functions of time. (*b*) To what form of energy is the kinetic energy of the puck converted?

7.4 Source energy

Because of the inexorable dissipation of mechanical energy, we need a source of energy—fuel, food, and so on—to either generate or maintain the mechanical energy of a system. For instance, the energy of gasoline is required to keep a car moving. Without this supply of energy, the car slows down as its kinetic energy dissipates because of friction. Likewise, to walk around the block, to send a message across the continent, to lift a basket, or to squeeze a spring, we need some source of energy.

Our current needs for mechanical energy are met using energy from a variety of sources—fossil and mineral fuels (oil, coal, natural gas), nuclear fuel, biomass fuel (food, wood, organic waste), water reservoirs, geothermal activity, solar radiation, and wind. We refer to the energy obtained from these sources collectively as **source energy.** Broadly speaking, there are four kinds of source energy: *chemical energy* (energy associated with the configuration of atoms inside molecules) released in such chemical reactions as the burning of oil, coal, gas, and wood and the metabolizing of food; *nuclear energy* (energy associated with the configuration of the nuclei of atoms) released in nuclear reactions; *solar energy* delivered by radiation from the Sun; and *stored solar energy* in the form of wind and hydroelectric energy.

Nearly all source energy can be traced back to energy delivered by the Sun. Nuclear reactions in the Sun create the solar radiation responsible for the energy content of biomass—the food we eat, the wood we burn. Fossil and mineral fuels were formed from the buried remains of plants and animals that lived and stored energy from the Sun millions of years ago. Wind energy results from the uneven heating of Earth's surface and atmosphere by solar radiation. Solar radiation evaporates the water that forms the clouds that deposit water in reservoirs behind dams.

7.7 How should chemical energy be classified in Figure 7.10?

CONCEPTS

To facilitate our accounting of energy in the rest of this book, we divide all energy into four categories: kinetic energy K, potential energy U, source energy E_s, and thermal energy E_{th} (**Figure 7.12a** on the next page). Interactions convert energy from one category to another, but conservation of energy requires that, for a closed system, the energy E of the system remains unchanged (Figure 7.12*b*).

Figure 7.13 on the next page illustrates the various types of energy conversions that can occur. Figure 7.13*a* shows the reversible conversion between the two forms of mechanical energy. Examples of this conversion are the (reversible) conversion between gravitational potential energy and kinetic energy for a ball thrown vertically upward and the (reversible) conversion between elastic potential energy and kinetic energy for a cart hitting a spring.

Figure 7.12 Energy accounting.

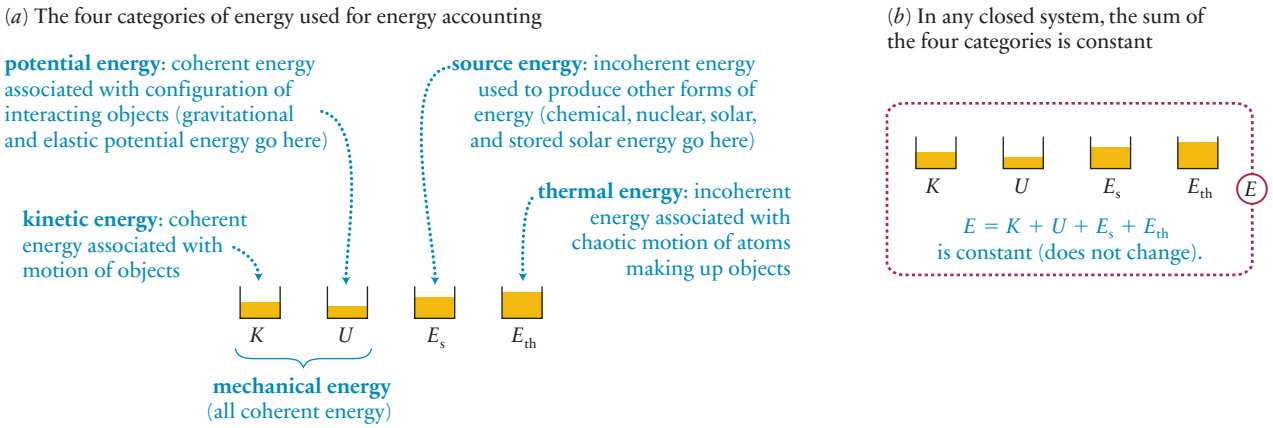


Figure 7.13*b* shows the (irreversible) dissipation of mechanical energy: Friction converts mechanical energy to thermal energy. Because of this dissipation, we need to convert source energy to maintain or generate mechanical energy (Figure 7.13*c*). For example, in a chemical reaction, the spatial arrangement of atoms inside molecules changes, and chemical energy is converted to kinetic energy of the reaction products (**Figure 7.14**). In the combustion of fuel in an engine, some of the kinetic energy of the combustion products can be converted to mechanical energy, but a large portion of it is converted to thermal energy (a car engine gets so hot that part of the mechanical energy generated by the engine must be used to turn a fan to *cool* the engine). Friction in the engine converts even more energy to thermal energy. Most of this thermal energy ends up either in the atmosphere or in the ground and cannot be easily recovered. We shall examine the generation of mechanical energy by *engines* in more detail in Chapter 21. For now, remember that the generation of mechanical energy from source energy is always accompanied by the generation of thermal energy, as illustrated in Figure 7.13*c*.

Figure 7.13 Energy conversion processes.

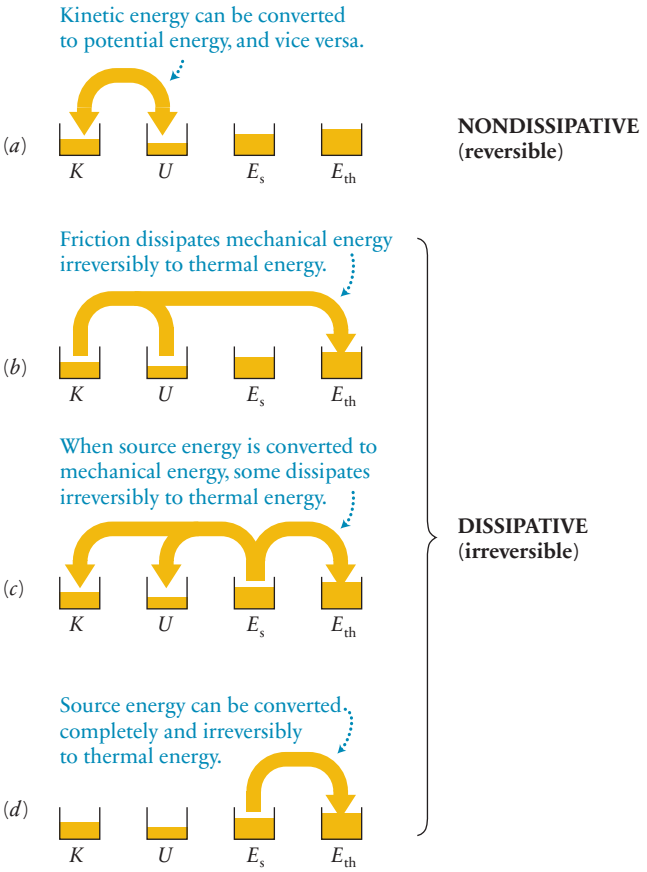
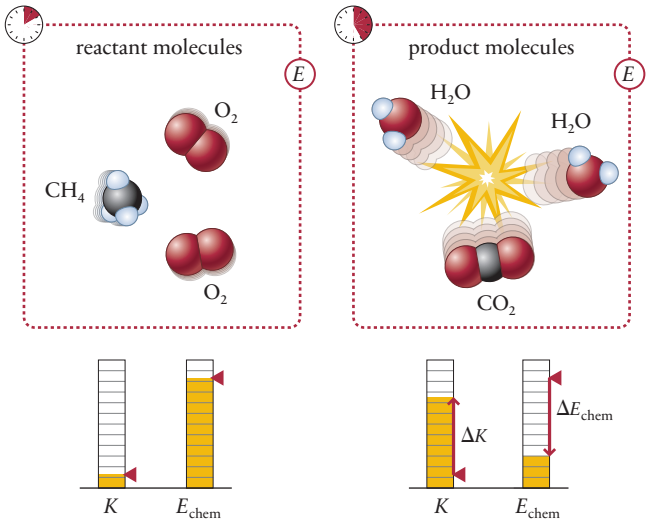


Figure 7.14 In the combustion of methane, each molecule of methane (CH_4) reacts with two oxygen molecules (O_2). The reaction rearranges the constituent atoms to form two water molecules (H_2O) and one molecule of carbon dioxide (CO_2). In the reaction, chemical energy originally in the reactant molecules is converted to kinetic energy of the products.



CONCEPTS

Figure 7.13*d* shows the conversion of source energy to thermal energy. An example of this process is the burning of gas or oil to heat our homes. The conversion of source energy to either mechanical or thermal energy is *irreversible*. The combustion of a methane-oxygen gas mixture, for example, causes permanent changes in the physical state of the gas mixture; it becomes a mixture of water and carbon dioxide. The reverse—obtaining methane from a mixture of water and carbon dioxide—is inconceivable even if energy is added to the system. Because of this irreversibility, the conversion of source energy to mechanical energy for transportation, heating, and so on—what we call *energy consumption*—decreases the available amount of source energy stored on Earth.

Because of dissipation, all forms of energy eventually are converted to thermal energy. As we shall see in Chapter 21, our ability to recover mechanical energy from thermal energy is very limited, however, making thermal energy not as useful as other forms of energy. Although some energy sources are renewable (solar, wind, water, biomass), they may not meet our future demands for energy. Consequently the supply of source energy is limited, and a number of sources, such as oil and gas, are likely to run out during the 21st century.

7.8 Whenever you leave your room, you diligently turn off the lights to "conserve energy." Your friend tells you that energy is conserved regardless of whether or not your lights are off. Which of you is right?

As we have seen, interactions that convert source energy to either mechanical or thermal energy and interactions that convert mechanical energy to thermal energy are irreversible. Irreversible interactions change the thermal energy in a system and are called **dissipative interactions;** reversible interactions do not change the thermal energy in a system and are called **nondissipative interactions.**

To determine whether or not an interaction is dissipative, all you need to do is ask yourself whether the interaction is reversible:

Interactions that cause reversible changes are nondissipative; those that cause irreversible changes are dissipative.

Consider a puck sliding to a stop on a rough surface. If you were to make a movie of this action and then play the movie in reverse, you would immediately notice that something is not correct: A puck lying on a

Coherent versus incoherent energy

To get a feel for the different types of energy, let's look at a very ordinary object—a pencil. How many types of energy are (or can be) stored in a pencil?

Coherent (mechanical) energy

Kinetic energy: If the pencil is at rest on your desk, its kinetic energy is zero. If, however, you throw it across the room, you give it a kinetic energy of about **0.01 J.**

Potential energy: To store potential energy in the pencil, you must change its configuration by squeezing or bending it. The pencil, being stiff, does not bend or squeeze easily, and so, if you are lucky, you might be able to store **0.1 J** of elastic potential energy in it before it snaps.

Incoherent energy

Thermal energy: Thermal energy is the energy associated with the random jiggling of atoms. It is impossible to convert all of this energy to another form, and so instead let's imagine you take the pencil out of a drawer and carry it around in your pocket so that its temperature rises from room temperature to body temperature. This rise in temperature increases its thermal energy by **100 J.** If you could convert this 100 J of thermal energy to kinetic energy, your pencil would be moving at the speed of an airplane!

Chemical energy: Being made of wood, a pencil stores chemical energy, which can easily be released. If you burn the pencil, the wood turns to ashes, and the configuration energy stored in the chemical bonds is converted to thermal energy, which you feel as heat. The energy converted by burning the pencil is **100,000 J,** an amount equal to the kinetic energy of a medium-sized car moving at 35 mi/h.

Broken chemical bonds: If you bend the pencil enough, it breaks. The energy required to break the chemical bonds in the pencil is about **0.001 J.** This amount is only 1% of that 0.1 J of potential energy you added by bending the pencil to just before its breaking point; the remaining 99% ends up as thermal energy: When the pencil snaps, the stress created by the bending is relieved, and the snapping increases the jiggling of the atoms.

This comparison of energies makes two important points that are valid more generally. First, coherent forms of energy are insignificant compared with incoherent forms—most of the energy around us is in the form of incoherent energy. Second, when energy is dissipated, virtually all of it becomes thermal energy; the incoherent configuration energy associated with deformation, breaking, and abrasion is generally negligible compared to the energies required to cause these changes.

rough surface doesn't spontaneously start moving. So the interaction between the rough surface and the puck is irreversible and therefore dissipative.

EXERCISE 7.3 Converting energy

For each of the following scenarios, choose a closed system, identify the energy conversions that take place, and classify each conversion according to the four processes shown in Figure 7.13. (*a*) A person lifts a suitcase. (*b*) A toy suspended from a spring bobs up and down. (*c*) A pan of water is brought to a boil on a propane burner. (*d*) A cyclist brakes and comes to a stop.

SOLUTION (*a*) Closed system: person, suitcase, and Earth. During the lifting, the potential energy of the Earth-suitcase system increases and the kinetic energy of the suitcase increases. The source energy is supplied by the person doing the lifting, who converts chemical (source) energy from food. In the process of converting this source energy, thermal energy is generated (the person gets hot). This process is represented in Figure 7.13*c*. ✔

(*b*) Closed system: toy, spring, and Earth. As the toy bobs, its height relative to the ground changes, its velocity changes, and the configuration of the spring changes. The bobbing thus involves conversions of gravitational and elastic potential energy and kinetic energy. This reversible process is represented in Figure 7.13*a*. ✔

(*c*) Closed system: pan of water and propane tank. As chemical (source) energy is released by burning the propane, the water is heated and its thermal energy increases. This process is represented in Figure 7.13*d*. ✔

(*d*) Closed system: Earth and cyclist. During the braking, the bicycle's kinetic energy is converted to thermal energy by friction (I'm ignoring the muscle source energy required to pull the brakes). This process is represented in Figure 7.13*b*. ✔

CONCEPTS



7.9 For each of the following processes, determine what energy conversion takes place and classify the interaction as dissipative or nondissipative. (Hint: Imagine what you would see if you played each situation in reverse.) (*a*) The launching of a ball by the expanding of a compressed spring, (*b*) the fall of a ball released a certain height above the ground, (*c*) the slowing down of a coasting bicycle, (*d*) the acceleration of a car.

7.5 Interaction range

The underlying mechanisms of interactions are still poorly understood, and as a result no one can answer the question, How do objects interact? From experience we know that not all objects interact with one another, but what determines when and why certain things do interact? It appears that each type of interaction occurs only between certain kinds of objects. So we assign to these objects certain *attributes* (magnetic strength, color, and so on) that are responsible for the interaction.

Note that it is the interaction that defines these attributes; we have no independent way of determining them. Consider a magnet, for instance. Is it a magnet because it behaves like one, or does it behave like a magnet because it *is* a magnet? In other words, do we say it has the attributes of a magnet because it interacts with other magnets? Or does it interact with other magnets because it has the attributes of a magnet?

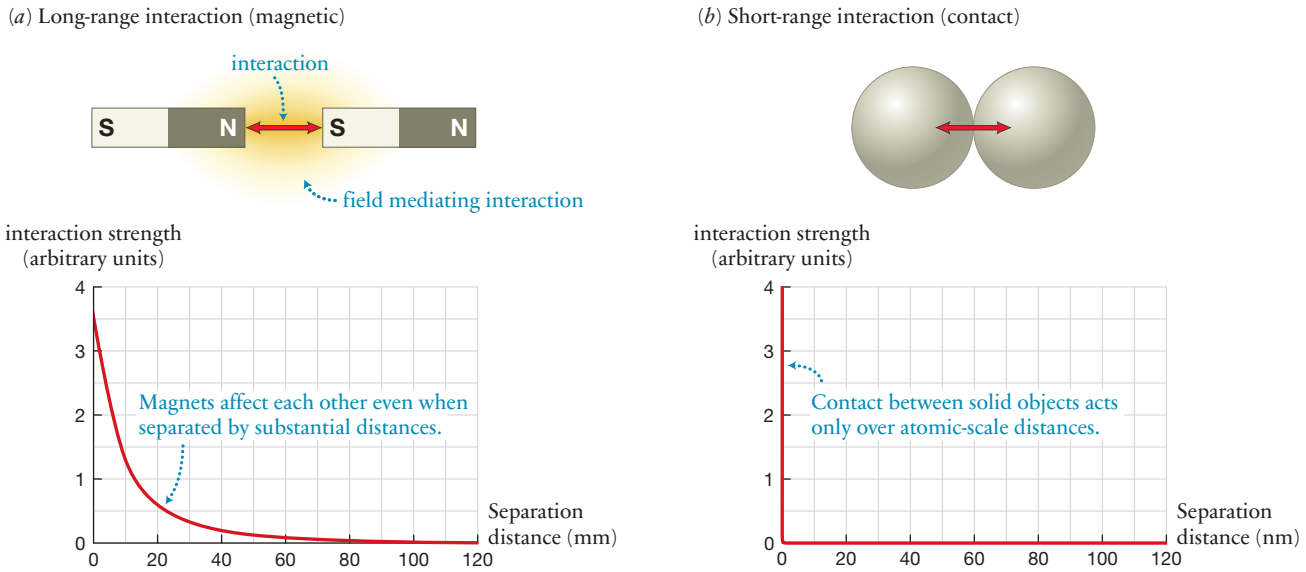
What all this boils down to is that we *define* the properties of matter (whether something is magnetic or nonmagnetic, heavy or light, red or green) according to the interactions the matter takes part in (it interacts with iron or doesn't, Earth pulls on it strongly or not, it interacts with green light or with red). Matter can therefore be classified according to its interactions, and the attributes we give to various types of matter are just a convenient way of indicating the types of interactions they take part in.

We are used to thinking of interactions as "things touching." You know, however, that a bar magnet can attract or repel another magnet located some distance away. The gravitational interaction works that way, too: Without any physical contact, Earth pulls downward on an object thrown up in the air. And so does electricity: On a day when the humidity is low, a plastic comb held near your head can attract your hair without physical contact.

To help us understand how things can interact without touching, we need to examine what we mean by *physical contact*. On the macroscopic level, *physical contact* means just that—two surfaces touching each other. On the atomic level, however, two atoms attract or repel each other even when they are separated by distances several times their size. Because all things are made up of atoms, there is no such thing as "touching" on the atomic level and therefore no such thing as physical contact. Your feet resting on the floor, for example, are supported by countless interactions between the atoms of your feet and those in the floor. So, on the atomic level, because atoms interact without physical contact, even feet "touching" floor is really an interaction without physical contact!

The strength of any interaction between two objects (in this context, think of *strength* in the sense of the magnitude of the acceleration caused by the interaction) is a function of the distance separating the objects. **Figure 7.15** shows how this strength varies with distance for magnets and for billiard balls. Because the magnets "feel" each other over quite a large distance, their interaction is said to be a *long-range* interaction.

Figure 7.15 Long-range interactions are perceptible over macroscopic distances; short-range ones, only over distances at the atomic scale or smaller. For instance, the contact interaction between two billiard balls is noticeable only when they are within nanometers of each other.



The repulsion (or, depending on the orientation of the magnets, attraction) is greatest at small separation distances but continues to be felt even at relatively large distances. In contrast, there is no noticeable interaction between the billiard balls even when they are separated by as little as the width of a hair. The two balls interact only when they are brought sufficiently close for the individual atoms to "feel" one another (which we call "touching"). The physical contact between two objects is thus a *short-range* interaction.

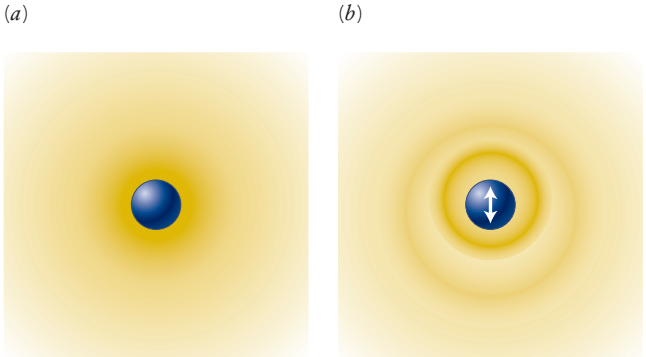
As these examples illustrate, our notion of physical contact has to do with how rapidly the effects of an interaction change with distance. The difference between the magnets and billiard balls in Figure 7.15 is their **interaction range**—the distance over which the interaction is appreciable.

One model widely used to picture long-range interactions is the **field** (**Figure 7.16***a*). In this model, every object is surrounded by a field, the properties of which are determined by the attributes of the object. An electrically charged object has an electric field, a magnetic object has a magnetic field, any object with what is called mass has a gravitational field, and so on. (We shall discuss mass in Chapter 13; for now just think of the mass of an object as a measure of how heavy that object is.) Interactions between objects are mediated by these fields.

Fields have an existence that is separate from the objects that create them. If I shake an object, the shaking creates a disturbance in the field that travels outward much like ripples on the surface of water (Figure 7.16*b*). The field of any object therefore contains information on the entire history of the object. As an example, think of watching the stars on a clear, moonless night. The light that reaches your eyes from some of the stars you are looking at was emitted as long as a million years ago. Some of the stars you see may not even exist anymore. Yet, via the ripples they once created in their fields—manifested as the light they once emitted—they are still interacting with you.

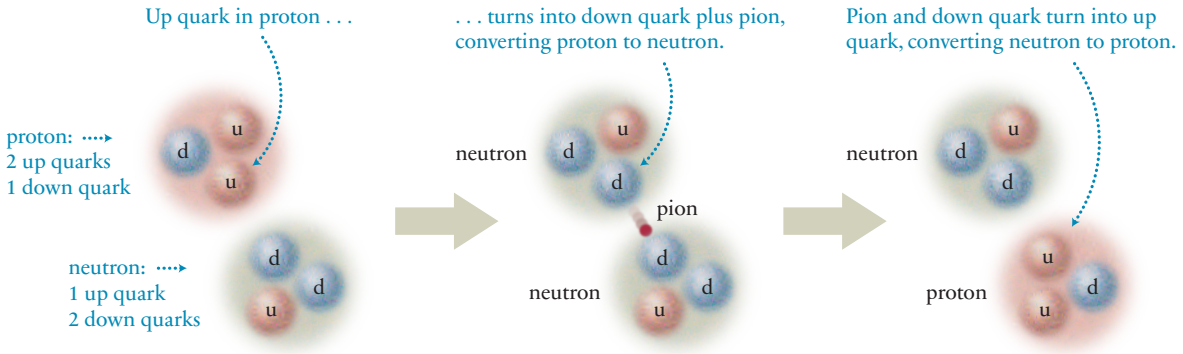
An alternative model explains interactions in terms of an exchange of fundamental particles called *gauge particles* (**Figure 7.17** and Checkpoint 7.10 on the next page). Physicists now believe that every interaction is caused by corresponding exchange of gauge particles. The existence of many of these particles has indeed been confirmed experimentally.

Figure 7.16 (*a*) The field of a stationary spherical object occupies the space around the object. The shading indicates the strength of the field—a measure of how strongly the object interacts with other objects. The field shown here rapidly decreases in strength with distance. (*b*) An outward traveling ripple in the field created by shaking the object.



CONCEPTS

Figure 7.17 Protons and neutrons (the particles that make up the atomic nucleus) can interconvert by exchanging a particle called a pion. As a result of this pion exchange, an example of which is shown schematically below, protons and neutrons exchange places continually as though "pulled through each other." The effect is that of an attractive interaction between protons and neutrons.



7.10 As an example of an interaction mediated by a "particle," imagine tossing a ball back and forth with a friend. You are both standing on an icy surface so slippery that friction is negligible. (*a*) Describe the effect the throwing and catching have on your momentum and your friend's momentum. (*b*) Is this an attractive or repulsive interaction? (*c*) Could you change the interaction into the other type?

7.6 Fundamental interactions

An interaction is **fundamental** if it cannot be explained in terms of other interactions. The interaction between, say, two colliding carts is not fundamental because it can be explained as being the result of the interactions between the atoms that make up the carts. Nor are the atomic interactions fundamental because they are the result of the interactions between the electrons and atomic nuclei that make up the atoms. And inside each atomic nucleus we observe yet another interaction responsible for holding the components of the nucleus together. This last interaction, the *strong interaction,* is today classified as fundamental because it cannot (yet) be explained in terms of other interactions.

All known interactions can be traced to just four fundamental interactions, two of them familiar from everyday life and two not (**Table 7.1**). The two familiar ones are the gravitational and electromagnetic interactions. We know of the gravitational interaction as Earth pulling on objects. A nylon sock interacts with a silk blouse in the clothes dryer via the electric part of the electromagnetic interaction; a magnet interacts with a refrigerator door via the magnetic part. The two unfamiliar fundamental interactions are those that take place inside the nuclei of atoms: the weak interaction and the strong interaction. A brief overview of the four fundamental interactions follows.

1. Gravitational interaction. This long-range interaction manifests itself as an attraction between all objects that have *mass,* and it is believed to be mediated by a still undetected gauge particle called the *graviton.* It is the weakest fundamental interaction, as manifested by the observation that you can't feel it acting between your body and any nearby object even though both of you possess mass. Only for very massive objects does this interaction become appreciable: The gravitational interaction between Earth and your body, for instance, *is* noticeable.

CONCEPTS

Table 7.1 Fundamental Interactions

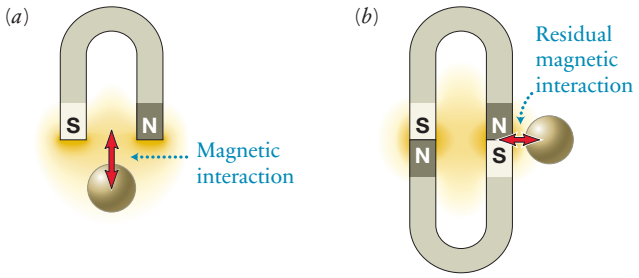
Type	Required attribute	Relative strength	Range	Gauge particle	Propagation speed
gravitational	mass	1	∞	graviton?	c?
weak	weak charge	10^{25}	10^{-18} m	vector bosons	varies
electromagnetic	electrical charge	10^{36}	∞	photon	c
strong	color charge	10^{38}	10^{-15} m	gluon	c

The relative strength is a measure of the magnitude of the effects of these interactions on two protons separated by about 10^{-15} m. The question marks in the last two columns indicate that the information provided has not yet been verified experimentally. The symbol c represents the speed at which lights travels.

Figure 7.18 The matter in a galaxy is held together by the gravitational interaction—an indication of the tremendous range of this interaction.



Figure 7.19 Residual interaction. Even though a single magnet interacts strongly with a piece of iron, a pair of magnets joined together interacts only weakly with it.



Even though it is weak and decreases with increasing distance between the interacting objects, the gravitational interaction is what determines the structure of the universe on the cosmic scale. Other interactions generally cease to play a role on this scale. The existence of planetary systems and the shape of galaxies, such as the one shown in **Figure 7.18**, indicate the enormous range of the gravitational interaction.

2. Electromagnetic interaction. It is no exaggeration to say that this long-range interaction is responsible for most of what happens around us. It is responsible for the structure of atoms and molecules, for all chemical and biological processes, for the cohesion of matter into liquids and solids, for the repulsive interaction between objects such as a bat and ball, as well as for light and other electromagnetic radiation. The attribute of matter responsible for this interaction is called *electrical charge,* which comes in two varieties: *positive* and *negative,* each appearing in equal numbers in the universe. The gauge particle associated with the electromagnetic interaction is the *photon* (discussed in Chapter 34).

For moving charged particles, the electromagnetic interaction is complex (we'll treat it later in this book). For charged particles at rest, the interaction is either attractive or repulsive, depending on the combination of charged particles involved. Like the gravitational interaction, the strength of the electromagnetic interaction decreases rapidly with increasing distance between the interacting objects. However, the electromagnetic interaction of two protons separated by about 10^{-15} m is 36 orders of magnitude (that is, 10^{36} times) stronger than the gravitational interaction. To get a feel for just how large this difference in magnitude is, consider that the inertia of a grain of sand is "only" 33 orders of magnitude less than the inertia of Earth.

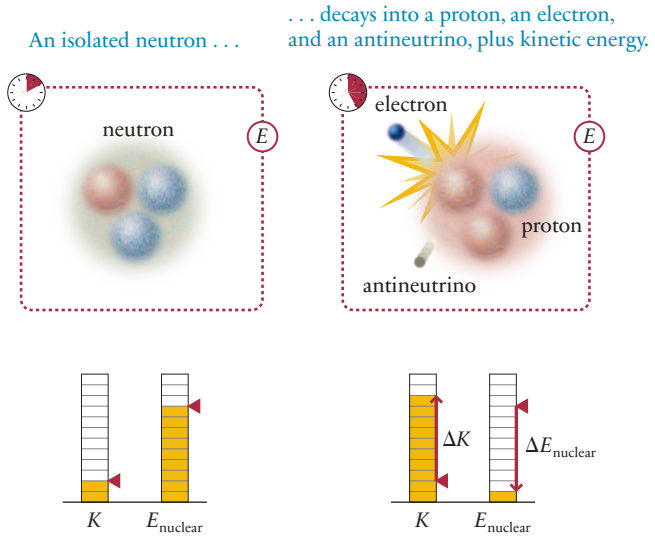
Because of the perfect balance in the amount of positive and negative charge in the universe and because the electromagnetic interaction is so strong, all matter tends to arrange itself in such a way that it is electrically neutral (which means it contains just as much positive as negative charge, and as a result most electromagnetic interactions cancel out). Consequently the electromagnetic interaction is most noticeable at the atomic scale. As noted in Section 7.5, what we call physical contact between objects is the result of the electromagnetic interactions between the individual charged parts of the atoms—interactions that average out at greater distances. The situation is much the same as when you let two horseshoe magnets attract each other until they stick together, as in **Figure 7.19**. The magnetic interaction between the pair and a piece of iron is considerably weaker than the interaction between that piece of iron and just one of the magnets. This is so because, in the pair, most of the effect of, say, the north pole of one magnet is canceled by the south pole of the other magnet. At short range, the canceling of the magnetic effect is not complete, and so there is some magnetic interaction—a *residual electromagnetic interaction*—left over, to interact weakly with the piece of iron.

3. Weak interaction. This (usually) repulsive interaction is responsible for some radioactive decay processes and for the conversion of hydrogen to helium in stars. The weak interaction acts inside the nucleus between subatomic particles that possess an attribute called *weak charge* and causes most subatomic particles to be unstable. The gauge particles associated with the weak interaction are called *vector bosons*.

There is a strong parallel between the role of the electromagnetic interaction in chemical reactions and the role of the weak interaction in the decay of subatomic particles. Chemical reactions are caused by the rearrangement of electrons among atoms or groups of atoms (Figure 7.14). In the process, chemical energy is converted to (or from) kinetic energy of the reaction products. Likewise, when the weak interaction causes a subatomic particle to decay, internal energy of the

CONCEPTS

Figure 7.20 A neutron that is not part of an atomic nucleus disintegrates into a proton, an electron, and a subatomic particle called the antineutrino. This disintegration, called beta decay, is caused by the weak interaction.



original particle is converted to kinetic energy of the decay products (**Figure 7.20**).

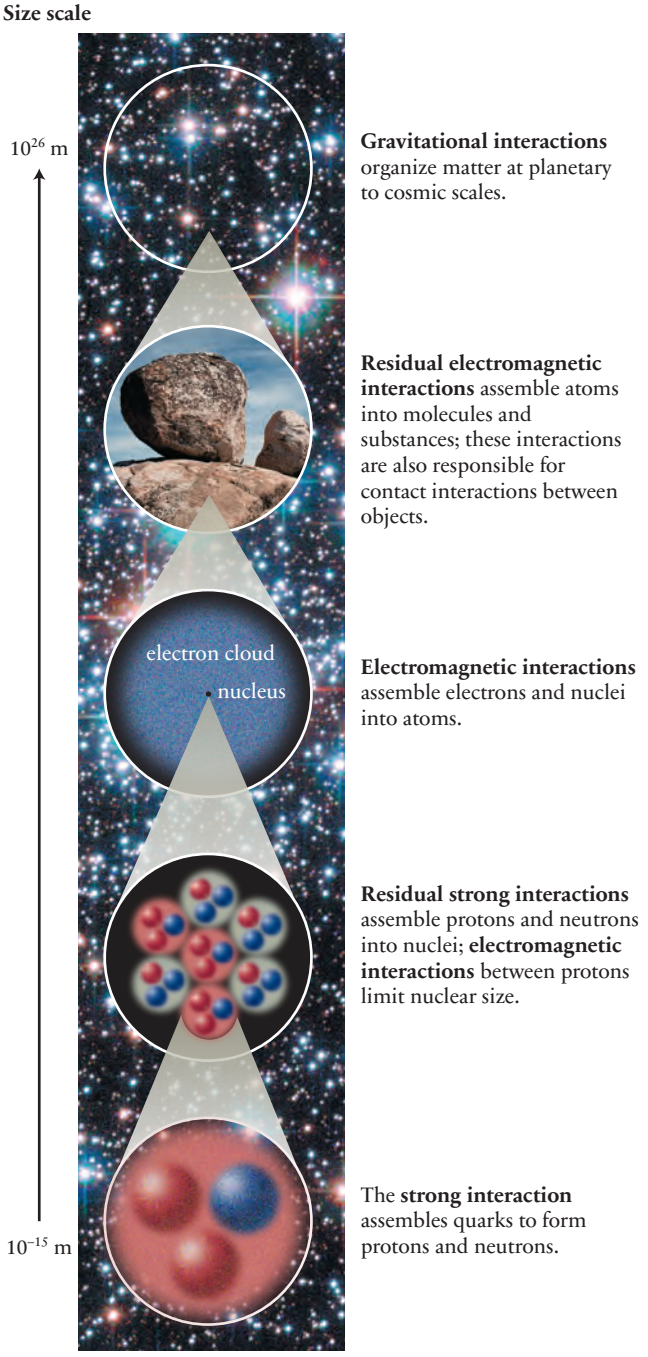
There is a strong connection between the electromagnetic and weak interactions. Both can be considered manifestations of a single interaction called the *electroweak interaction*.

4. Strong interaction. The strong interaction acts between *quarks*, which are the building blocks of protons, neutrons, and other particles. This interaction, which can be either attractive or repulsive, is so strong that it completely overwhelms all other interactions among particles. The attribute required for this interaction is *color charge*, and the gauge particles are called *gluons*.

Quarks, like electrically charged particles and magnets, always arrange themselves in such a way that the effects of the strong interaction cancel. Only at very short range does some of the strong interaction remain. This *residual strong interaction* is responsible for holding the nucleus of an atom together.

In Table 7.1 note that the gravitational interaction—the fundamental interaction we are most directly aware of in our daily life—is orders of magnitude weaker than the other fundamental interactions. This is so because the stronger the interaction, the smaller the scale on which it affects the structure of the universe (**Figure 7.21**). (The size scale shown in the figure is not linear or logarithmic.)

Figure 7.21 The fundamental interactions determine the structure of the universe at different scales. At the subatomic scale, the strong interaction is responsible for fusing quarks into the protons and neutrons that make up the atomic nucleus. The residual strong interaction and the weak interaction, along with electromagnetic repulsion between protons, are responsible for the formation and stability of atomic nuclei. Outside the atomic nucleus all effects of the strong interaction cancel, and weaker interactions become noticeable. Electromagnetic interactions between the positively charged atomic nucleus and negatively charged electrons are responsible for the structure of atoms. Interactions between atoms are due to residual electromagnetic interactions. At the human scale, the picture is dominated by the short-range residual electromagnetic interactions involved in contact between objects and by gravitational interactions. At the cosmic scale, the gravitational interaction plays the main role.



CONCEPTS

Self-quiz

1. Two carts are about to collide head-on on a track. The inertia of cart 1 is greater than the inertia of cart 2, and the collision is elastic. The speed of cart 1 before the collision is higher than the speed of cart 2 before the collision. (*a*) Which cart experiences the greater acceleration during the collision? (*b*) Which cart has the greater change in momentum due to the collision? (*c*) Which cart has the greater change in kinetic energy during the collision?
2. Which of the following deformations are reversible and which are irreversible: (*a*) the deformation of a tennis ball against a racquet, (*b*) the deformation of a car fender during a traffic accident, (*c*) the deformation of a balloon as it is blown up, (*d*) the deformation of fresh snow as you walk through it?
3. Translate the kinetic energy graph in Figure 7.2 into three sets of energy bars: before the collision, during the collision, and after the collision. In each set, include a bar for K_1, a bar for K_2, and a bar for the internal energy of the system, and assume that the system is closed.
4. Describe a scenario to fit the energy bars shown in **Figure 7.22**. What happens during the interaction?

Figure 7.22

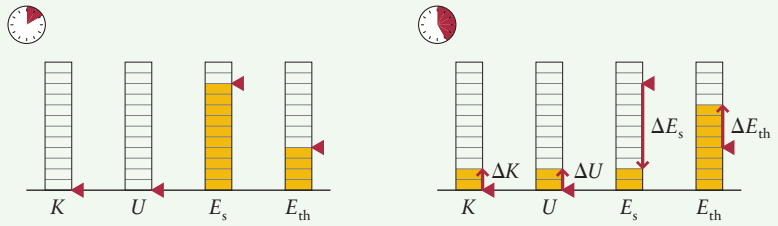
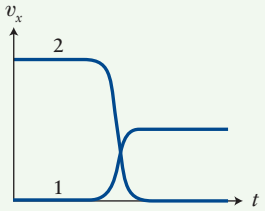


Figure 7.23

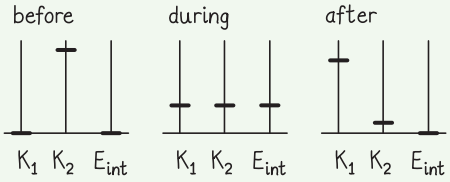


5. Describe a scenario to fit the velocity-versus-time curves for two colliding objects shown in **Figure 7.23**. All motion is confined to the x axis. What happens to the initial energy of the system of colliding objects during the interaction?

Answers

1. (*a*) The cart with the smaller inertia experiences the greater acceleration (see Figure 7.2). (*b*) The magnitude of $\Delta\vec{p}_1$ is the same as the magnitude of $\Delta\vec{p}_2$, but the changes are in opposite directions because the momentum of the system does not change during the collision. (*c*) $|\Delta K_1| = |\Delta K_2|$, but the changes are opposite in sign because the kinetic energy of the system before the elastic collision has to be the same as the kinetic energy of the system afterward.
2. (*a*) Reversible. The ball returns to its original shape. (*b*) Irreversible. The fender remains crumpled. (*c*) Irreversible. The balloon does not completely return to its original shape after deflation. (*d*) Irreversible. Your footprints remain.
3. See **Figure 7.24**. Before the collision $K_1 = 0$, K_2 is maximal, and $E_{int} = 0$; during the collision K_1, K_2, and E_{int} are all about one-third of the initial value of K_1; after the collision K_1 is about 7/8 of the initial value of K_1, K_2 is about 1/8 of the initial value of K_1, and $E_{int} = 0$. Because the system is closed, its energy is constant, which means the sum of the three bars is always the same.

Figure 7.24



4. During the interaction, eight units of source energy is converted to two units of kinetic energy, two units of potential energy, and four units of thermal energy. One possible scenario is the vertical launching of a ball. Consider the system comprising you, the ball, and Earth from just before the ball is launched until after it has traveled some distance upward: The source energy goes down (you exert some effort), thermal energy goes up (in the process of exerting effort you heat up), kinetic energy goes up (the ball was at rest before the launch), and so does potential energy (the distance between the ground and the ball increases).
5. The graph represents an inelastic collision because the relative velocity of the two objects decreases to about half its initial value. In order for the momentum of the system to remain constant, the inertia of object 1 must be twice that of object 2. Possible scenario: Object 2, inertia m, collides inelastically with object 1, inertia $2m$. The collision brings object 2 to rest and sets object 1 in motion. The interaction converts the initial kinetic energy of object 2 to kinetic energy of cart 1 and to thermal energy and/or incoherent configuration energy of both carts.

7.7 Interactions and accelerations

Let us now prove a statement made in Section 7.1: For interactions that affect the motion of objects, the ratio of the x components of the accelerations of the interacting objects is equal to the negative inverse of the ratio of their inertias.

Whenever two objects interact, they can exchange momentum and energy. If we consider the two-object system to be isolated during the interaction, conservation of momentum requires that, at any instant during the interaction, the momentum changes of the two objects be equal in magnitude and opposite in direction:

$$\Delta\vec{p}_1 = -\Delta\vec{p}_2. \tag{7.1}$$

Because the change in momentum for each object is caused by the interaction with the other object, the time interval over which the momenta change—the duration of the interaction—is the same for both objects. If we denote this time interval by Δt, we can write

$$\frac{\Delta\vec{p}_1}{\Delta t} = -\frac{\Delta\vec{p}_2}{\Delta t}. \tag{7.2}$$

If the inertias m_1 and m_2 of the two objects are not changed by the interaction, we have, from the definition $\vec{p} \equiv m\vec{v}$,

$$\frac{m_1\Delta\vec{v}_1}{\Delta t} = -\frac{m_2\Delta\vec{v}_2}{\Delta t}. \tag{7.3}$$

Using the definition of acceleration given as Eq. 3.21,

$$\lim_{\Delta t\to 0}\frac{\Delta v_x}{\Delta t} \equiv a_x, \tag{7.4}$$

we take the limit of both sides of Eq. 7.3 as Δt goes to zero and get for the x components of the accelerations:

$$m_1a_{1x} = -m_2a_{2x}. \tag{7.5}$$

Hence, both objects accelerate during the interaction, and the ratio of the x components of their accelerations is equal to the negative inverse of the ratio of their inertias:

$$\frac{a_{1x}}{a_{2x}} = -\frac{m_2}{m_1}. \tag{7.6}$$

This relationship between the accelerations of two interacting objects of constant inertia holds for all interactions in an isolated two-object system, regardless of what happens to the energy during the interaction.

QUANTITATIVE TOOLS



7.11 A 1000-kg compact car and a 2000-kg van, each traveling at 25 m/s, collide head-on and remain locked together after the collision, which lasts 0.20 s. (*a*) According to Eq. 7.6, their accelerations during the collision are unequal. How can this be if both initially have the same speed and the time interval during which the collision takes place is the same amount of time for both? (*b*) Calculate the average acceleration in the direction of travel experienced by each vehicle during the collision.

7.8 Nondissipative interactions

Conservation of energy requires that the energy of any closed system remain unchanged. Using the four categories of energy introduced in Section 7.4, we can write this in the form

$$\Delta E = \Delta K + \Delta U + \Delta E_s + \Delta E_{th} = 0 \quad \text{(closed system)}. \tag{7.7}$$

For nondissipative interactions, there are no changes in either the source energy or the thermal energy of the system: $\Delta E_s = 0$ and $\Delta E_{th} = 0$. Thus the only energy conversions allowed in nondissipative interactions are reversible transformations between kinetic energy K and potential energy U. For these types of interactions Eq. 7.7 simplifies to

$$\Delta E = \Delta K + \Delta U = 0 \quad \text{(closed system, nondissipative interaction)}. \tag{7.8}$$

If we introduce the mechanical energy of a system

$$E_{mech} = K + U, \tag{7.9}$$

we can write Eq. 7.8 as

$$\Delta E_{mech} = 0 \quad \text{(closed system, nondissipative interaction)}. \tag{7.10}$$

In words, in a closed system in which only nondissipative interactions take place, the mechanical energy (the sum of the kinetic and potential energies) is constant.

We can deduce an important property of potential energy from the reversibility of nondissipative interactions. Consider the nondissipative interaction shown in **Figure 7.25*a*** on the next page. A cart, initially traveling to the right, bounces off a spring. The cart interacts with the spring, which is anchored to Earth. The system comprising the cart, spring, track, and Earth is closed. The interaction between the cart and Earth (which we'll consider to include the track and the spring) is nondissipative because it is reversible: Reversing the sequence yields the same result. From the motion of the cart, you cannot tell if time runs from t_1 to t_5 or the other way around (Figure 7.25*b*).

This reversibility means that the cart's kinetic energy at some given position x must be the same on the way in and on the way out. For example, K_{cart} at t_2—when the front end of the cart is at x_2—must be the same as K_{cart} at t_4—when the front end returns to x_2. Because K_{Earth} doesn't change (see Checkpoint 7.12), Eq. 7.8 yields

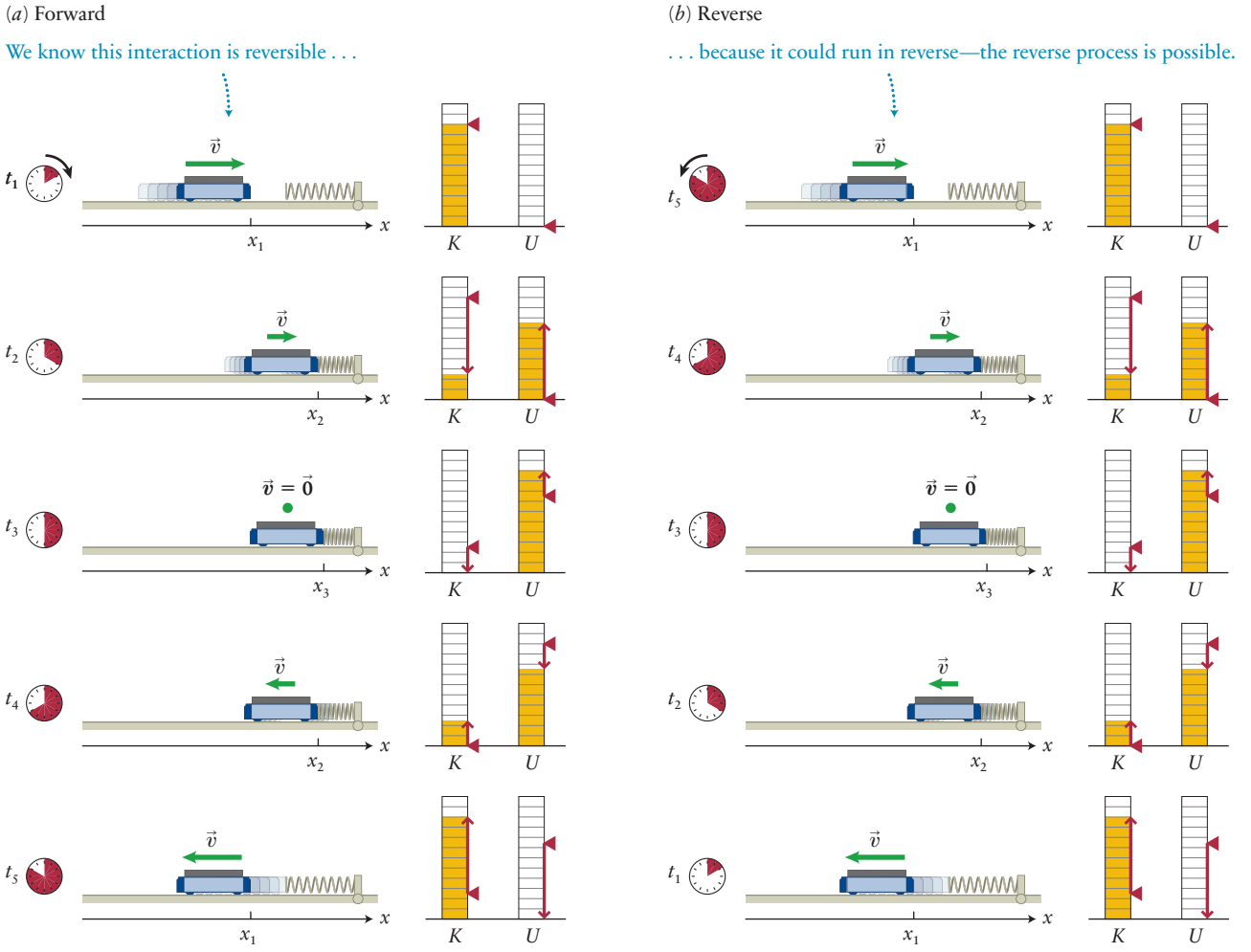
$$\Delta K_{cart} = -\Delta U_{spring}, \tag{7.11}$$

where U_{spring} is the elastic potential energy associated with the shape of the spring. Consequently, the potential energy in the spring must also have the same value at any given position on the way in and on the way out. In other words, the potential energy stored in the spring must be a unique function of position: U_{spring} has a definite value at each position x of the end of the spring. More generally, the potential energy of any system (that is, the portion of the system's internal energy that is associated with reversible configuration changes) can be written in the form

$$U = U(x), \tag{7.12}$$

where $U(x)$ is a unique function of a position variable x that quantifies the configuration of the system.

Figure 7.25 Reversible interaction between a cart and a spring anchored to a post.



A direct consequence of this dependence of potential energy on position is that the change in the kinetic energy of an object that moves from a position x_1 to a position x_2 during some nondissipative interaction depends *only* on those positions, not on the path taken by the object. Example 7.4 shows how this applies to the specific case of a cart bouncing off a spring.

7.12 Use the conservation laws to show that, when the spring in Figure 7.25 expands, the change in the kinetic energy of Earth is negligible and we are therefore justified in using Eq. 7.11.

EXAMPLE 7.4 Path independence of change in potential energy

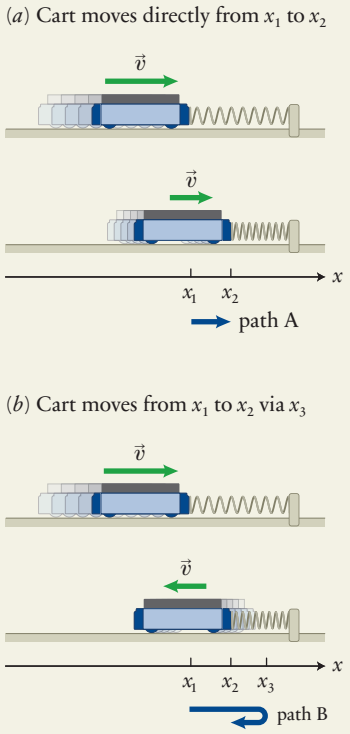
Figure 7.26 shows a cart striking a spring. In Figure 7.26*a*, consider the motion of the cart along the direct path from the initial position x_1, which is the position at which the cart makes initial contact with the free end of the spring, to the position x_2 (path A). In Figure 7.26*b*, consider the motion along the path from x_1 to the position of maximum compression x_3 and then back to x_2 (path B). Show that the change in the cart's kinetic energy is the same for both paths if the interaction caused by the spring is reversible.

❶ **GETTING STARTED** As the cart moves from x_1 to x_2 along path A, the spring is compressed and the cart's kinetic energy steadily decreases. Along path B, the spring is first compressed and the cart comes to a stop at x_3. The spring then expands, accelerating the cart back to x_2. To solve this problem, I'll consider the closed system made up of the cart, the spring, and Earth.

❷ **DEVISE PLAN** The change in the cart's kinetic energy is given by Eq. 7.11. Therefore if the change in the elastic

QUANTITATIVE TOOLS

Figure 7.26 Example 7.4.



potential energy of the spring is the same along both paths, then the change in the kinetic energy of the cart must also be the same along both paths. I'll first determine the change in potential energy between the initial and final positions of path A, then do the same between the initial and final positions of the two parts of path B.

3 EXECUTE PLAN Because the potential energy associated with a reversible interaction is a function of x only (Eq. 7.12), the change in elastic potential energy along path A is

$$\Delta U_{\text{path A}} = U_f - U_i = U(x_2) - U(x_1).$$

Along path B I have two contributions to ΔU_{spring}: an increase in U from x_1 to x_3 and a decrease from x_3 to x_2:

$$\begin{aligned}\Delta U_{\text{path B}} &= \Delta U_{13} + \Delta U_{32} \\ &= [U(x_3) - U(x_1)] + [U(x_2) - U(x_3)] \\ &= U(x_2) - U(x_1) = \Delta U_{\text{path A}}. \checkmark\end{aligned}$$

So, the change in the cart's kinetic energy is the same for both paths.

4 EVALUATE RESULT The change in the cart's kinetic energy is independent of the path joining x_1 and x_2 even though the cart ends up moving in opposite directions along the two paths. This is because the change in potential energy ΔU_{12} depends only on the coordinates of the initial and final positions x_1 and x_2.

Another way to see that this must be true is to consider only the motion from x_2 to x_3 and back. From x_2 to x_3 all the kinetic energy of the cart is converted to elastic potential energy of the spring. From x_3 to x_2 the elastic potential energy is converted back to kinetic energy, and so the motion from x_2 to x_3 and back to x_2 doesn't change the cart's kinetic energy. (Remember that kinetic energy is always positive and independent of the direction of travel.)

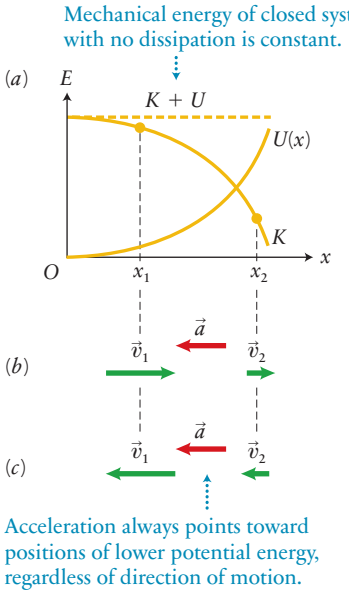
7.13 Show that, in Figure 7.26, the change in potential energy along a round trip from position x_1 to position x_2 and then back to x_1 is zero.

We can deduce a very important point from the fact that potential energy is a unique function of position. Consider, for example, a closed system with the potential energy function $U(x)$ shown in **Figure 7.27a**. If there is no dissipation, the sum of the kinetic and potential energies of this closed system is constant. Therefore, because the potential energy increases with increasing x, the kinetic energy must decrease with increasing x. So, if we consider two positions x_1 and $x_2 > x_1$, we see that the kinetic energy must be smaller at x_2 than at x_1. If the object whose kinetic energy is represented in the graph moves toward increasing x, as shown in Figure 7.27*b*, it slows down, and so its acceleration must point to the left. If, on the other hand, the object moves toward smaller x, as shown in Figure 7.27*c*, it speeds up, and so its acceleration again points to the left. This means that the object tends to be accelerated in the direction of lower potential energy, regardless of its original direction of motion. Quite generally we can state:

The parts of any closed system always tend to accelerate in the direction that lowers the system's potential energy.

In Figure 7.26, for example, the potential energy increases as the free end of the spring moves to the right. The free end therefore accelerates the

Figure 7.27 Acceleration in a closed system that has no energy dissipation.



cart to the left. This acceleration slows the cart down as it moves to the right and speeds it up as it moves to the left.

7.14 Consider a ball launched upward. Verify that its acceleration points in the direction that lowers the gravitational potential energy of the Earth-ball system.

7.9 Potential energy near Earth's surface

As a first example of how to determine the potential energy for a nondissipative interaction, consider a familiar interaction: the gravitational interaction near Earth's surface. As we discussed in Chapter 3, it is an experimental fact that, in the absence of air resistance, all objects near the ground fall with constant acceleration $g = 9.8\ \text{m/s}^2$. If we ignore friction, this interaction is nondissipative because free fall is reversible (see Checkpoint 7.9*b*).

If we release a ball from a certain height above the ground, the ball gains kinetic energy as it accelerates downward. As the ball falls, the distance between it and the ground decreases so that the (gravitational) configuration of the Earth-ball system and hence the potential energy associated with this configuration change. Because the ball and Earth constitute a closed system and because the gravitational interaction is nondissipative, we can use Eq. 7.8 to relate the change in the ball's kinetic energy to a change in the gravitational potential energy U^G (as we found in Checkpoint 7.12, Earth's kinetic energy doesn't change):*

$$\Delta U^G + \Delta K_b = 0, \tag{7.13}$$

where ΔU^G is the change in the gravitational potential energy of the Earth-ball system and ΔK_b is the change in the ball's kinetic energy.

Equation 7.13 allows us to obtain ΔU^G if we know ΔK_b. Using some of the kinematics from Chapter 3, we can obtain an expression that gives ΔU^G in terms of physical quantities other than the kinetic energy. Let's begin by looking at our ball as it falls from x_i to x_f in a time interval $\Delta t = t_f - t_i$. As illustrated in **Figure 7.28**, at t_i the ball is at a height x_i above the ground and the x component of its velocity is $v_{x,i}$; at t_f the ball is at x_f and the x component of its velocity is $v_{x,f}$. In that time interval, the ball accelerates with a constant acceleration given by Eqs. 3.1 and 3.14:

$$a_x = \frac{\Delta v_x}{\Delta t} = -g. \tag{7.14}$$

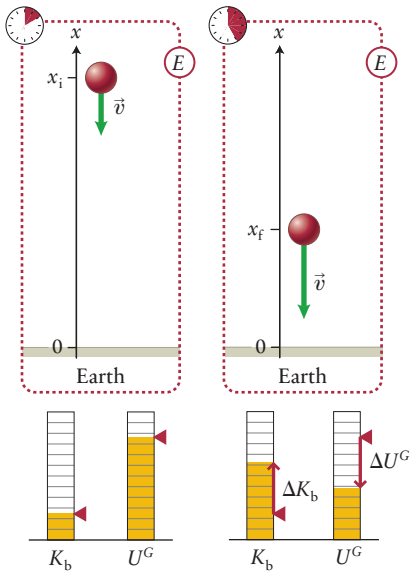
Next we turn to Eq. 3.7, which when we substitute $-g$ for a_x is

$$\Delta x = v_{x,i}\Delta t - \tfrac{1}{2}g(\Delta t)^2. \tag{7.15}$$

Rearranging Eq. 7.14 gives $\Delta t = -\Delta v_x/g$, and substituting this expression for Δt into Eq. 7.15 gives us

$$\Delta x = -v_{x,i}\frac{\Delta v_x}{g} - \tfrac{1}{2}g\left(\frac{\Delta v_x}{-g}\right)^2. \tag{7.16}$$

Figure 7.28 A freely falling ball gains kinetic energy at the expense of the gravitational potential energy of the (closed) Earth-ball system.



*We use superscripts to indicate the type of interaction associated with the potential energy: G for gravitational, E for electric, and so on.

QUANTITATIVE TOOLS

Expanding the delta notations and working through the algebra, we get

$$x_f - x_i = -\frac{2v_{x,i}(v_{x,f} - v_{x,i}) + v_{x,f}^2 - 2v_{x,i}v_{x,f} + v_{x,i}^2}{2g} = -\frac{v_f^2 - v_i^2}{2g}, \qquad (7.17)$$

where we have used the fact that $v_x^2 = v^2$. Multiplying through by g and by the inertia m_b of the ball, we get

$$m_b g(x_f - x_i) + \tfrac{1}{2}m_b(v_f^2 - v_i^2) = 0. \qquad (7.18)$$

The second term on the left is precisely the change in the ball's kinetic energy, and so comparison with Eq. 7.13 tells us that the first term must be ΔU^G, the change in gravitational potential energy we are looking for. When the distance between the ball and the ground changes by an amount Δx, the gravitational potential energy changes by an amount $m_b g \Delta x$.

Because *all* objects falling near Earth's surface experience the same acceleration, this reasoning holds for any object. So, if the vertical coordinate of any object of inertia m changes by Δx, the change in the gravitational potential energy of the Earth-object system is

$$\Delta U^G = mg\Delta x \quad \text{(near Earth's surface)}. \qquad (7.19)$$



7.15 Suppose you raise this book (inertia $m = 3.4$ kg) from the floor to your desk, 1.0 m above the floor. (*a*) Does the gravitational potential energy of the Earth-book system increase or decrease? (*b*) By how much? (*c*) Conservation of energy requires that this change in potential energy be compensated for by a change in energy somewhere in the universe. Where?

Equation 7.19 gives the change in gravitational potential energy between two positions:

$$\Delta U^G = U_f^G - U_i^G = mg(x_f - x_i) = mgx_f - mgx_i, \qquad (7.20)$$

from which we deduce that the gravitational potential energy of the Earth-object system is given by

$$U^G(x) = mgx \quad \text{(near Earth's surface)}, \qquad (7.21)$$

where x is the object's vertical coordinate with the positive x axis pointing up.

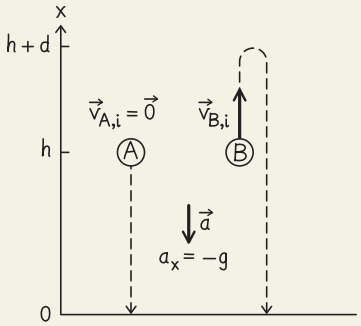
Implicit in this expression is that $U^G = 0$ at $x = 0$. Note, however, that if we write $U^G = mg(x + c)$, with c some arbitrary constant, Eq. 7.20 remains the same because the constant c, which appears in both terms, cancels out. This arbitrariness in the reference point for zero gravitational potential energy is no different from the arbitrariness in the reference point for measuring elevations or from the reference frame dependence of kinetic energy—physically, only *changes* in energy are relevant. Because of this arbitrariness, you can choose whichever reference point you like as the position where the gravitational potential energy is zero (for example, the ground at $x = 0$, as in Eq. 7.21).

EXAMPLE 7.5 Path independence of gravitational potential energy

Ball A is released from rest at a height h above the ground. Ball B is launched upward from the same height at initial speed $v_{B,i}$. The two balls have the same inertia m. Consider this motion from the instant they are released to the instant they hit the ground. (*a*) Using kinematics, show that the change in kinetic energy is the same for both balls. (*b*) Show that this change in kinetic energy is equal to the negative of the change in the gravitational potential energy of the Earth-ball system between positions $x = +h$ and $x = 0$.

❶ **GETTING STARTED** I begin by representing these two motions graphically (**Figure 7.29**). Ball A drops straight down; ball B first rises a distance d above the launch position h and then falls from a height $h + d$.

Figure 7.29



❷ **DEVISE PLAN** I can obtain the initial kinetic energies from the initial speeds, which are given. So all I need to do is calculate the final speeds using the kinematics equations for motion at constant acceleration from Chapter 3. From my sketch, I see that the x component of the acceleration of the balls is given by $a_x = -g$. For part *b* I can use Eq. 7.20 to calculate the change in gravitational potential energy.

❸ **EXECUTE PLAN** (*a*) The time interval ball A takes to fall to the ground can be obtained from Eq. 3.7:

$$\Delta x = x_f - x_i = 0 - h = v_{Ax,i}\Delta t - \tfrac{1}{2}g(\Delta t)^2$$

$$= -\tfrac{1}{2}g(\Delta t)^2$$

$$(\Delta t)^2 = \frac{2h}{g}. \qquad (1)$$

The x component of the ball's final velocity is, from Eq. 7.14, $v_{Ax,f} = v_{Ax,i} - g\Delta t = -g\Delta t$, and so, using Eq. 1, I calculate that the kinetic energy of ball A just before it strikes the ground is

$$K_{A,f} = \tfrac{1}{2}mv_{A,f}^2 = \tfrac{1}{2}m[-g\Delta t]^2 = \tfrac{1}{2}mg^2\left(\frac{2h}{g}\right) = mgh. \qquad (2)$$

Because the ball begins at rest, $K_{A,i} = 0$, and so the change in kinetic energy is $\Delta K_A = K_{A,f} - K_{A,i} = mgh$.

Equation 2 holds for any value of h. After reaching its highest position, ball B falls a distance $h + d$ to the ground, and so, substituting $h + d$ for h in Eq. 2, I obtain the final kinetic energy of ball B: $K_{B,f} = mg(h + d)$. To determine d, I examine ball B's motion from its initial position $x_i = h$ to the top of its path, where $x_f = h + d$, and again use Eq. 3.7:

$$\Delta x = (h + d) - h = d = v_{Bx,i}\Delta t_{top} - \tfrac{1}{2}g(\Delta t_{top})^2, \qquad (3)$$

but the time interval Δt_{top} required to reach the top of the trajectory is not known. I do know, however, that the ball's velocity at the top is zero, and so I can say, again using Eq. 7.14, that $v_{Bx,top} = 0 = v_{Bx,i} - g\Delta t_{top}$, or $\Delta t_{top} = v_{Bx,i}/g$. Equation 3 then becomes

$$d = v_{Bx,i}\left(\frac{v_{Bx,i}}{g}\right) - \tfrac{1}{2}g\left(\frac{v_{Bx,i}}{g}\right)^2 = \frac{v_{B,i}^2}{g} - \frac{1}{2}\frac{v_{B,i}^2}{g} = \frac{1}{2}\frac{v_{B,i}^2}{g}.$$

The final kinetic energy of ball B is thus

$$K_{B,f} = mg(h + d) = mgh + mg\left(\frac{1}{2}\frac{v_{B,i}^2}{g}\right) = mgh + \tfrac{1}{2}mv_{B,i}^2.$$

Because $K_{B,i} = \frac{1}{2}mv_{B,i}^2$ is the initial kinetic energy of ball B, the gain in kinetic energy for ball B is $\Delta K_b = K_{B,f} - K_{B,i} = mgh$, which is indeed equal to that of ball A (Eq. 2). ✔

(*b*) According to Eq. 7.20, the change in gravitational potential energy is

$$\Delta U^G = mg(x_f - x_i) = mg(0 - h) = -mgh. \checkmark$$

This result is equal in magnitude to the gains in the balls' kinetic energies I found in part *a*.

❹ **EVALUATE RESULT** Because I obtained the same result in two ways, I can be quite confident of my answer. My result further confirms that the change in potential energy is independent of the path taken and depends on only the endpoints of the path.

7.16 Suppose that instead of choosing Earth and the ball as our system in the discussion leading up to Eq. 7.21, we had chosen to consider just the ball. Does it make sense to speak about the gravitational potential energy of the ball (the way we speak of its kinetic energy)?

QUANTITATIVE TOOLS

7.10 Dissipative interactions

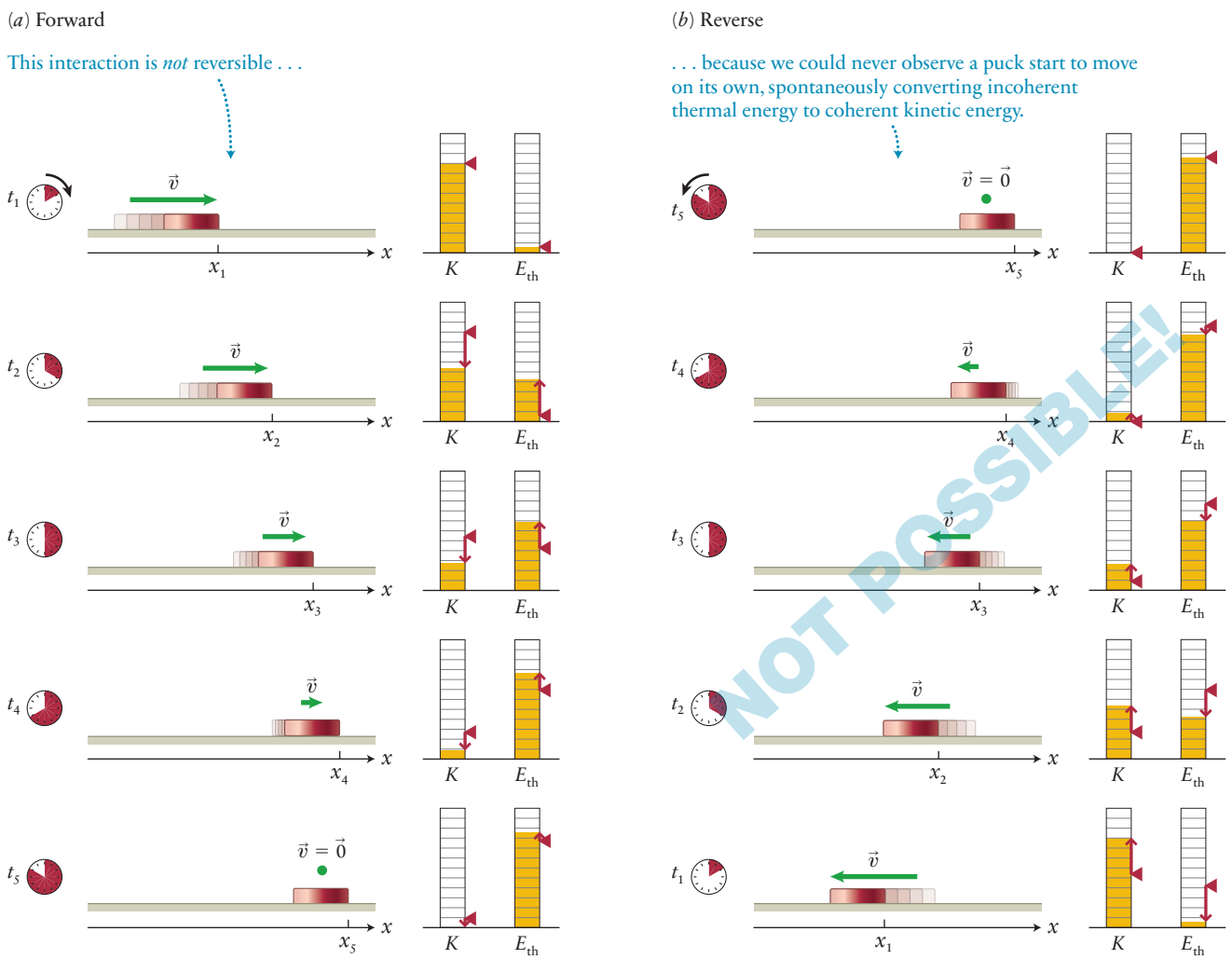
As noted in Sections 7.3 and 7.4, a dissipative interaction involves changes in thermal energy, source energy, or both. These interactions so improbable that we say they are *irreversible*. Consider **Figure 7.30*a***, which shows a puck sliding along a wooden board and slowing to a stop. The friction between the board and the puck is dissipative: If we play the sequence in reverse, we obtain an impossible situation (Figure 7.30*b*). The friction acting on the puck in Figure 7.30*a* causes its kinetic energy to be converted to thermal energy—to be dissipated. The system comprising the puck, the board, and Earth (which holds the board in place) is closed, and so Eq. 7.7, $\Delta E = \Delta K + \Delta U + \Delta E_s + \Delta E_{th} = 0$, holds. We know that ΔU is zero because there is no change in the puck's height relative to Earth and there are no springs involved. The source energy has to be zero because no nuclear or chemical reactions are involved, and so Eq. 7.7 simplifies to

$$\Delta K + \Delta E_{th} = 0. \tag{7.22}$$

Because the only kinetic energy that changes is that of the puck (the changes in the kinetic energies of the board and Earth are negligible, see Checkpoint 7.12), we can say

$$\Delta K_{puck} = -\Delta E_{th}. \tag{7.23}$$

Figure 7.30 A puck that slides to a halt (*a*) dissipates energy through friction. Because this interaction is dissipative, the process is irreversible; the sequence shown in (*b*) cannot occur.



Although this expression looks similar to Eq. 7.11, the right side is very different conceptually. In particular, thermal energy is not a function of position the way U is. The motion shown in Figure 7.30 can take place *anywhere* along the board (unlike the motion shown in Figure 7.25, which can take place only when the cart is at the spring).

Another way to see that the interaction between the puck and the board is dissipative is to look at what happens during a round trip. Imagine putting the puck down at some position x_A and then moving it back and forth between x_A and some other position x_B a dozen times. When the puck eventually returns to x_A, its state is changed: It is hotter because of the friction (and some wear has taken place). This means that energy has been dissipated irreversibly, and so $\Delta E_{th} \neq 0$ even though the puck is back at its starting position. Moving the cart of Figure 7.25 back and forth a dozen times, on the other hand, yields no change in the spring's potential energy: $\Delta U_{spring} = 0$ whenever the cart is back at its starting position.

Equation 7.7 also describes the changes in energy that take place during inelastic collisions. We know that ΔU is zero because any reversible configuration changes that occur during the first half of the collision are undone during the second half. Provided no source energy is converted during the collision, conservation of energy applied to the closed system of the colliding objects requires that the changes in kinetic energy and thermal energy add up to zero, as stated in Eq. 7.22, which we can restate in the form

$$\Delta K = -\Delta E_{th}. \tag{7.24}$$

This equation is applicable to elastic, inelastic, and totally inelastic collisions. In elastic collisions, no energy is dissipated. Thus $\Delta E_{th} = 0$ and so Eq. 7.24 reduces to

$$\Delta K = 0 \quad (\text{elastic collision}), \tag{7.25}$$

which means that the kinetic energy of the system does not change, as we already know from Eq. 5.15. For inelastic collisions, we can compare Eq. 7.24 and Eq. 6.43, $\Delta K = \frac{1}{2}\mu v_{12i}^2 (e^2 - 1)$, where the reduced inertia μ is $m_1 m_2/(m_1 + m_2)$ and the coefficient of restitution e (Eq. 5.18) is v_{12f}/v_{12i}. This comparison tells us that the amount of kinetic energy converted to thermal energy is determined by the coefficient of restitution:

$$\Delta E_{th} = -\Delta K = \tfrac{1}{2}\mu v_{12i}^2 (1 - e^2). \tag{7.26}$$

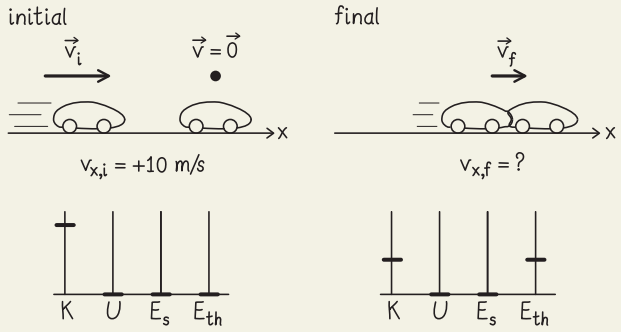
The smaller the value of e, the larger the amount of energy dissipated from coherent energy to incoherent energy.

EXAMPLE 7.6 Fender bender

A 1000-kg car traveling at 10 m/s collides with and attaches to an identical car that is initially at rest. (*a*) How much energy is dissipated in the collision? (*b*) Use Eq. 7.26 to verify your result.

❶ **GETTING STARTED** The collision is totally inelastic, and so the kinetic energy of the two-car system after the collision is less than it was before the collision. (It has to be nonzero, however, to maintain the momentum of the isolated two-car system.) Because the system can be considered closed during the collision and because no energy is stored reversibly ($\Delta U = 0$), the kinetic energy that disappears is converted to thermal energy (the crumpling heats the metal). I make the sketch in **Figure 7.31** to represent the energy changes in the system.

Figure 7.31



QUANTITATIVE TOOLS

2 DEVISE PLAN Equation 7.24 tells me that the increase in thermal energy in the system is equal to the decrease in kinetic energy. I can determine the initial kinetic energy from the data given. To obtain the final kinetic energy, I first use conservation of momentum to calculate the velocity of the two cars immediately after impact and then calculate the kinetic energy of the combined cars.

3 EXECUTE PLAN (*a*) The initial kinetic energy is $\frac{1}{2}(1000\text{ kg}) \times (10\text{ m/s})^2 = 50\text{ kJ}$. After the collision, the cars, locked together and forming an object of inertia 2000 kg, must have the same combined momentum as the single car before the collision, and so

$$(2000\text{ kg})v_{x,\text{f}} = (1000\text{ kg})(+10\text{ m/s})$$
$$v_{x,\text{f}} = +5.0\text{ m/s}.$$

The final kinetic energy is thus

$$\tfrac{1}{2}(2000\text{ kg})(5.0\text{ m/s})^2 = 25\text{ kJ}$$

so $$\Delta K = 25\text{ kJ} - 50\text{ kJ} = -25\text{ kJ}.$$

The amount of energy dissipated is 25 kJ. ✔

(*b*) For a totally inelastic collision, $e = 0$ and Eq. 7.26 becomes

$$\Delta E_{\text{th}} = \tfrac{1}{2}\mu v_{12\text{i}}^2 = \tfrac{1}{2}\left(\frac{m_1 m_2}{m_1 + m_2}\right)v_{12\text{i}}^2 = \frac{m^2}{4m}v_{12\text{i}}^2 = \tfrac{1}{4}mv_{12\text{i}}^2,$$

where I have used the fact that the two cars have the same inertia m. The initial relative speed $v_{12\text{i}}$ is equal to the initial speed of car 1, and so $E_{\text{th}} = \frac{1}{4}(1000\text{ kg})(10\text{ m/s})^2 = 25\text{ kJ}$, as I found before. ✔

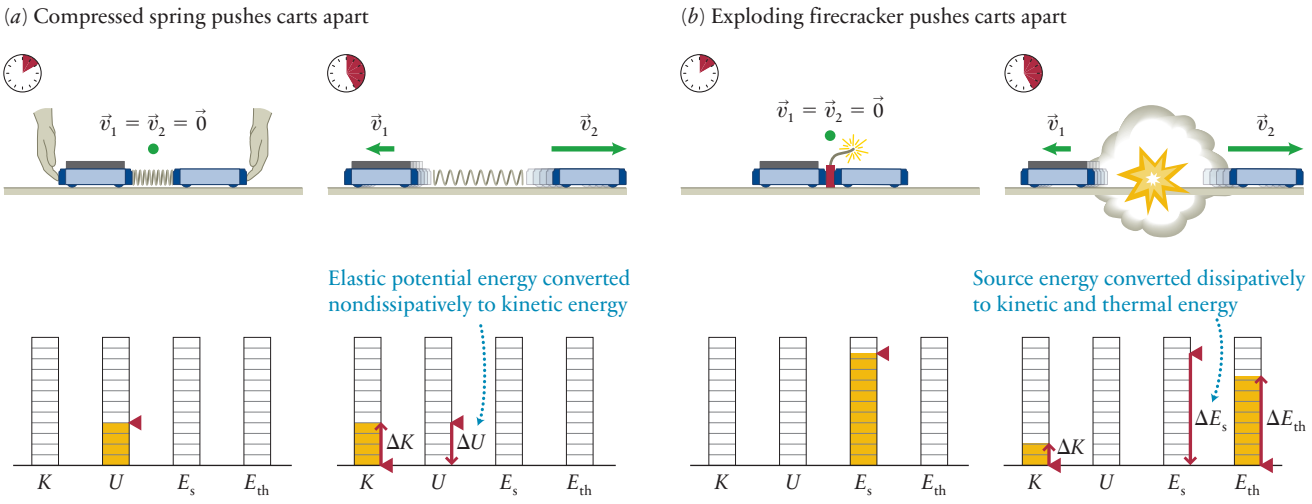
4 EVALUATE RESULT The two methods yield the same result, which is reassuring. The energy dissipated is equal to the chemical energy stored in just one-fourth of a wooden pencil (see the box on page 221), and yet both cars are likely to be wrecked beyond repair. Although it is surprising that one could wreck two cars with the chemical energy stored in one pencil, the box on page 221 states that coherent forms of energy are insignificant compared with incoherent forms, leading me to believe that the answer I obtained is not unreasonable.

As we saw in Section 5.8, a system's kinetic energy increases in an explosive separation. **Figure 7.32** shows two explosive separations. In Figure 7.32*a* the increase in kinetic energy is due to energy stored in a spring: One form of coherent energy (the elastic potential energy in the spring) is converted to another form of coherent energy (the kinetic energy of the moving carts). Because the compression of the spring is reversible, this interaction is nondissipative, and so this type of explosive separation is described by Eq. 7.8, $\Delta K + \Delta U = 0$.

The situation in Figure 7.32*b* is irreversible. Two carts are placed back to back against a firecracker, and when the firecracker is ignited, the explosive separation accelerates the carts away from each other. There is no change in potential energy because there is no change in the cart's height relative to Earth and there are no springs involved. There is, however, a change in source energy: During the explosive separation, chemical energy is converted partly to (coherent) kinetic energy of the carts and partly to (incoherent) thermal energy (things get hot during an explosion). For this dissipative interaction, we have, with $\Delta E_{\text{s}} = \Delta E_{\text{chem}}$,

$$\Delta K + \Delta E_{\text{chem}} + \Delta E_{\text{th}} = 0. \qquad (7.27)$$

Figure 7.32 Interactions that increase the kinetic energy of a system (*a*) nondissipatively and (*b*) dissipatively.



More generally, for any interaction involving a rise in temperature, or a permanent deformation in one or more of the interacting objects, or the conversion of source energy—in short, for any dissipative interaction—we have

$$\Delta K + \Delta U + \Delta E_s + \Delta E_{th} = 0 \quad \text{(closed system, dissipative interaction)}. \quad (7.28)$$

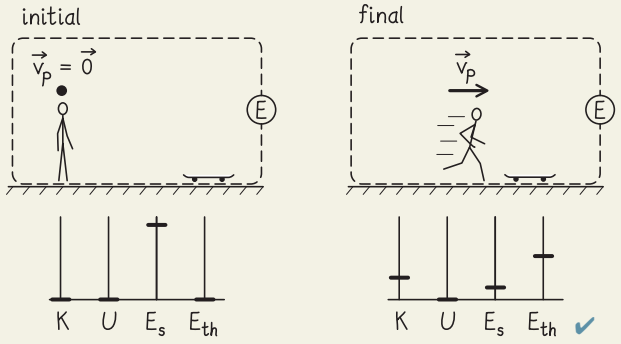
In general $\Delta E_s \leq 0$ because source energy is converted to another form of energy and $\Delta E_{th} \geq 0$ because dissipation increases the amount of thermal energy.

EXERCISE 7.7 Skateboard

Standing a few meters from a skateboard, a boarder begins running and then jumps onto the board. Make bar diagrams like the ones in Figure 7.32 to illustrate the energy conversions (*a*) while he accelerates and (*b*) while he jumps onto the board. Are the interactions dissipative or nondissipative?

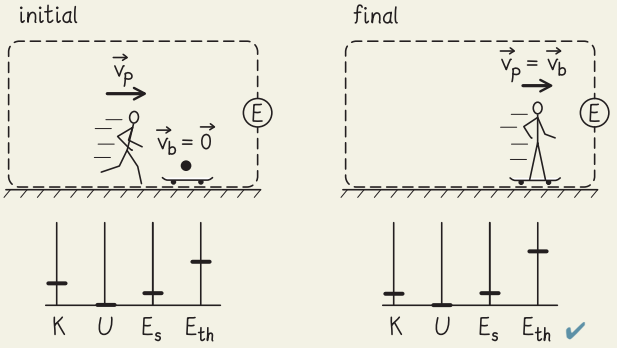
SOLUTION (*a*) I choose a closed system consisting of the boarder, the skateboard, and Earth. The boarder's initial kinetic energy is zero. As he accelerates during his run, he gains kinetic energy while the kinetic energy of Earth and the board remain zero. As he accelerates, he converts source energy (chemical energy extracted from food) to kinetic energy (he speeds up) and thermal energy (he gets warmer). This conversion is irreversible and therefore dissipative. There are no reversible changes in potential energy, and therefore I draw the energy bars shown in **Figure 7.33**.

Figure 7.33



(*b*) The jump onto the skateboard is a totally inelastic collision. Initially the boarder moves while the skateboard is at rest; after he lands on the board, both move at the same velocity. During this "collision," there is no conversion of source energy and no change of potential energy. As in any other totally inelastic collision, however, part of the initial kinetic energy is dissipated to thermal energy between the boarder's feet and the board and between the wheels and the ground. My bar diagram is shown in **Figure 7.34**.

Figure 7.34



7.17 How much energy is dissipated in the collision of Checkpoint 7.11?

Chapter Summary

The basics of interactions (Sections 7.1, 7.5, 7.6, 7.7)

Concepts An **interaction** is an event that produces either a physical change or a change in motion. A *repulsive interaction* causes the interacting objects to accelerate away from each other, and an *attractive interaction* causes them to accelerate toward each other.

The **interaction range** is the distance over which an interaction is appreciable. A long-range interaction has an infinite range; a short-range interaction has a finite range.

A **field** is a model used to visualize interactions between objects. According to this model, each object that takes part in an interaction produces a field in the space surrounding itself, and the fields mediate the interaction between the objects.

A **fundamental interaction** is one that cannot be explained in terms of other interactions. The four known fundamental interactions are the **gravitational interaction** (a long-range attractive interaction between objects that have *mass*), the **electromagnetic interaction** (a long-range interaction between objects that have *electrical charge;* this interaction can be either attractive or repulsive), the **weak interaction** (a short-range repulsive interaction between subatomic particles), and the **strong interaction** (a short-range interaction between *quarks,* the building blocks of protons, neutrons, and certain other subatomic particles; this interaction can be either attractive or repulsive).

Quantitative Tools If two objects of inertias m_1 and m_2 interact, the ratio of the x components of their accelerations is

$$\frac{a_{1x}}{a_{2x}} = -\frac{m_2}{m_1}. \tag{7.6}$$

Potential energy (Sections 7.2, 7.8, 7.9)

Concepts **Potential energy** is a coherent form of internal energy associated with reversible changes in the *configuration* of an object or system. Potential energy can be converted entirely to kinetic energy.

Gravitational potential energy is the potential energy associated with the relative position of objects that are interacting gravitationally.

Elastic potential energy is the potential energy associated with the reversible deformation of objects.

Changes in potential energy are *independent of path*. This means that the change in an object's potential energy as the object moves from a position x_1 to any other position x_2 depends *only* on x_1 and x_2, and *not* on the path the object takes in moving from x_1 to x_2.

Quantitative Tools The potential energy U of a system of two interacting objects can always be written in the form

$$U = U(x), \tag{7.12}$$

where $U(x)$ is a unique function of a position variable x that quantifies the configuration of the system.

Near Earth's surface, if the vertical coordinate of an object of inertia m changes by Δx, the gravitational potential energy U^G of the Earth-object system changes by

$$\Delta U^G = mg\Delta x. \tag{7.19}$$

Energy dissipation during interactions (Sections 7.3, 7.4, 7.8, 7.10)

Concepts All energy can be divided into two fundamental classes: energy associated with motion (kinetic energy) and energy associated with the configuration of interacting objects (potential energy). Each class of energy comes in two forms: *coherent* and *incoherent*. Energy is coherent if it involves ordered motion or configuration; it is incoherent if it involves random motion or configuration. For example, the kinetic energy of a moving object is coherent because all of its atoms move in the same way, whereas the thermal energy of an object is incoherent because the atoms move randomly.

The **mechanical energy** E_{mech} of a system is the sum of the system's coherent energies (kinetic energy K and potential energy U).

The **thermal energy** E_{th} of a system is the incoherent internal energy associated with the random motion of atoms that make up the objects in the system. Thermal energy cannot be entirely converted to coherent energy.

Source energy E_s is incoherent energy (such as chemical, nuclear, solar, and stored solar energy) used to produce other forms of energy.

Dissipative interactions are irreversible interactions that involve changes in thermal energy.

Nondissipative interactions are reversible interactions that convert kinetic energy to potential energy, and vice versa.

Quantitative Tools

$$E_{\text{mech}} = K + U. \tag{7.9}$$

During a dissipative interaction, the sum of the changes in all forms of energy in a closed system is zero:

$$\Delta K + \Delta U + \Delta E_s + \Delta E_{\text{th}} = 0. \tag{7.28}$$

During a nondissipative interaction, the mechanical energy of a closed system does not change:

$$\Delta E_{\text{mech}} = \Delta K + \Delta U = 0. \tag{7.8}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

7.1 The effects of interactions

1. A piece of salami is squashed between two pieces of bread. Is there one interaction between the salami and bread or more than one? •
2. Explain why you cannot cause Earth to accelerate very much, even if you jump up and down on it. •
3. Sketch curves to scale of momentum for the collision of object A initially moving with velocity v and object B initially at rest, for the following cases: (*a*) $m_A = m_B$, (*b*) $m_A = 2m_B$, (*c*) $m_A \gg m_B$. ••
4. Two objects of inertias m_1 and m_2 start from rest and then interact with each other (assume neither is interacting with any other object). (*a*) What is the ratio of their x components of velocity at any instant? (*b*) What is the ratio of their kinetic energies at any instant? ••
5. Two toy cars ($m_1 = 0.200$ kg and $m_2 = 0.250$ kg) are held together rear to rear with a compressed spring between them. When they are released, the cars are free to roll away from the ends of the spring. If you measure the acceleration of the 0.200-kg car to be 2.25 m/s^2 to the right while the spring is expanding, at that moment, what is the acceleration of the other car? ••

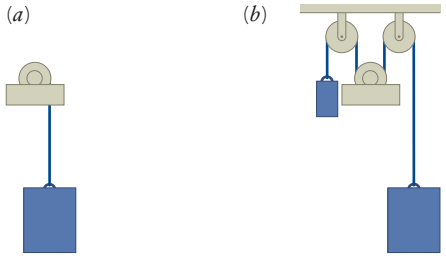
7.2 Potential Energy

6. You are at the swimming pool and want to make a splash. You climb the ladder to the high diving board, take a big bounce, and produce a prodigious splash. Describe the changes in kinetic and potential energy involved in this motion, beginning when you start to climb and ending when you hit the water. •
7. A 1-kg block is used to compress a spring with spring constant $k = 10.0$ N/m by 5 cm, and then released. Draw energy diagrams for the spring-mass system

when the spring is (*a*) fully compressed and (*b*) fully relaxed. •

8. A ring is attached at the center of the underside of a trampoline. A sneaky teenager crawls under the trampoline and uses the ring to pull the trampoline slowly down while his 75-kg mother is sleeping on it. When he releases the trampoline, she is launched upward. As she passes through the position at which she was before her son stretched the trampoline, her speed is 3.0 m/s. How much elastic potential energy did the son add to the trampoline by pulling it down? Assume the interaction is nondissipative. ••
9. An elevator might be raised and lowered directly by a winch, as in Figure P7.9*a*, or the mechanism might include a counterweight that moves up when the elevator moves down, as in Figure P7.9*b*. In which case is the change in the gravitational potential energy of the Earth-elevator-winch system lower? ••

Figure P7.9



10. Figures 7.7 and 7.25 suggest that potential energy in a nondissipative interaction builds up and then diminishes. Is this always the case? ••
11. Two blocks are held together with a compressed spring between them on the surface of a slippery table. One block has three times the inertia of the other. When the blocks are released, the spring pushes them away from each other. What is the ratio of their kinetic energies after the release? •••

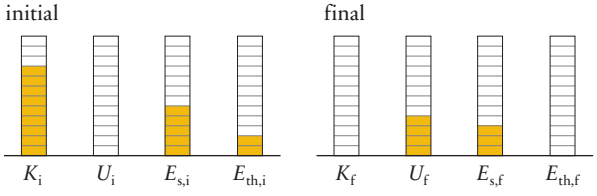
7.3 Energy dissipation

12. A trampoline's springs are what allow you to bounce, and if you bounce long enough, the springs get warm. Using Figure 7.10 as a guide, classify the energy in the springs that accounts for (*a*) the bouncing and (*b*) the heating up. (*c*) Where does this energy come from? •
13. You throw a baseball straight up, and because of air resistance some of the ball's energy is converted to thermal energy during the whole trip. Which takes longer: the trip up or the trip down? •
14. An arrow is shot into a hollow pipe resting on a horizontal table and flies out the other end. While the arrow travels in the pipe, its feathers brush against the walls of the pipe. (*a*) Which type of collision is the arrow-pipe interaction: elastic, inelastic, or totally inelastic? (*b*) Is there an instant when the velocity of the arrow relative to the pipe is necessarily zero? (*c*) Describe the energy conversions in the pipe-arrow system. ••
15. Imagine that you take a metal spring, compress it to its most compact length, and tie some string around it to keep it compressed. There is now potential energy stored in the compressed spring. You then place the tied-up spring into a container of strong acid, and, after some time, the compressed spring completely dissolves (the string is not affected by the acid). What type of energy does the compressed spring have and what is it converted to when it dissolves in the acid? ••
16. Mattresses are often a linked set of coil springs. Using Figure 7.10 as a guide, classify the energy that you give the mattress in each of these cases: (*a*) You slide the mattress from one side of a room to the other side. (*b*) You jump on the mattress several times. (*c*) You run your fingers along the tops of the coils, so that they jiggle around. (*d*) By jumping hard, you manage to snap some of the links between the coils. ••
17. Make a graph of the kinetic energy, potential energy, and incoherent energy versus time for (*a*) a moving billiard ball colliding with a stationary billiard ball, and (*b*) a moving ball of cookie dough colliding with, and sticking to, an identical but stationary ball of cookie dough. ••
18. When a steel ball and a soft rubber ball are dropped from the same height onto a concrete surface, the steel ball bounces higher. When they are dropped onto packed sand, however, the rubber ball bounces higher. Explain the difference. •••

7.4 Source energy

19. Choose an appropriate closed system and draw a bar diagram representing the energy conversions and transfers that occur during each process of Checkpoint 7.9: (*a*) a ball launching as the compressed spring it sits on expands, (*b*) a ball released from some height and falling to the ground, (*c*) a coasting bicycle slowing down, (*d*) a car accelerating on a highway. •
20. A 2000-kg car burns gasoline with 25% efficiency. Accelerating from rest, how fast would this car go upon burning 0.040 L of gasoline if 1.0 L of gasoline contains approximately 3.2×10^7 J of energy? ••
21. (*a*) Figure P7.21 shows initial and final energy bars for a closed system. Adjust the height of any two energy bars in the final version, on the right, so that the diagram does not violate what we know about energy. (There are numerous ways for you to do this.) (*b*) Describe an interaction that would generate an energy bar diagram that matches your version of Figure P7.21. ••

Figure P7.21



22. Two objects initially moving at velocities that have x components v_{1x} and v_{2x} collide totally inelastically. (*a*) Show that the ratio of the kinetic energy dissipated in the interaction is

$$\frac{\Delta K_1}{\Delta K_2} = -\frac{v_{\text{cm}x} + v_{1x,\text{i}}}{v_{\text{cm}x} + v_{2x,\text{i}}}.$$

(*b*) What does the minus sign in the expression in part *a* mean? ●●

7.5 Interaction range

23. If, on the atomic scale, "contact" between two objects is in fact not *physical* contact, does that have implications for the contact between atoms in a piece of copper? ●
24. Look again at Checkpoint 7.10. Suppose you and a friend stand within an arm's length of each other and throw a ball back and forth. Describe a method to model an attractive interaction based on throwing the ball. ●●
25. Look again at Checkpoint 7.10. (*a*) When you throw a ball to a friend, at what instant does your momentum change: at the beginning, middle, or end of the ball's flight? (*b*) At which of these three instants does your friend's momentum change? ●●
26. Give three examples to support the statement "long-range interaction forms the underlying basis for most life processes and technology." ●●

7.6 Fundamental interactions

27. Suppose that, instead of two types of electrically charged particles in the universe, there was only one type. What would you expect to happen to the structure of the universe if the particles produced (*a*) an attractive interaction and (*b*) a repulsive interaction? ●
28. Explain why tension (in ropes, etc.) and the contact forces between surfaces are not fundamental forces. ●
29. Suppose that the range of the strong interaction were suddenly increased by 20 orders of magnitude, with all other interactions unchanged. Describe the changes that might occur in the structure of our world. ●●
30. Picture two objects, A and B, taking part in some interaction involving gauge particles. If object A radiates gauge particles uniformly in all directions, how does the strength of the interaction depend on the distance from A to B? Assume that the gauge particles do not interact with one another and that the strength of the interaction is proportional to the number of gauge particles encountered by B. ●●●

7.7 Interactions and accelerations

31. On a low-friction track, a 1.7-kg cart initially going at 2.8 m/s to the right collides with a cart of unknown mass initially going at 1.2 m/s to the left. After the collision, both of the carts move to the right with the 1.7-kg cart at 2.0 m/s and the cart of unknown mass at 5.9 m/s. The collision takes 0.050 s. (*a*) Calculate the unknown mass. (*b*) Calculate the average acceleration of each cart. (*c*) Use your answer to part *a* to check your answer to part *b*. ●
32. Two children on ice skates pull toward each other on a rope held taut between them. The inertia of one child is 30 kg, and the inertia of the other is 25 kg. (*a*) If at one instant the 30-kg child is accelerating at 1.0 m/s^2 to the left, what is the acceleration of the 25-kg child at that instant? (*b*) What happens to each acceleration if the 25-kg child stops pulling and just hangs on? ●
33. A 0.010-kg bullet is fired from a 5.0-kg gun with a muzzle velocity of 250 m/s. (*a*) While the bullet is traveling in the barrel, what is the ratio of the acceleration of the gun to the acceleration of the bullet as a function of time? (*b*) What is the speed of the bullet relative to the ground at the instant the bullet leaves the barrel? (*c*) What is the speed of the gun relative to the ground at that instant? ●●
34. A 1500-kg car going at 6.32 m/s collides with a 3000-kg truck at rest. If the collision is totally inelastic and takes place over an interval of 0.203 s, what is the average acceleration of each vehicle? Are the ratio of accelerations and the ratio of inertias related the way you expect them to be? ●●
35. A glob of putty thrown to the right bounces off a 0.500-kg cart initially at rest on a low-friction track. The collision takes 0.15 s, and the coefficient of restitution is 0.64. The cart has a final velocity of 1.0 m/s to the right, and the glob has a final velocity of 0.834 m/s to the left. (*a*) What is the inertia of the glob? (*b*) What is the average acceleration of each object during the collision? ●●
36. An 8.20-kg object is sliding across the ice at 2.34 m/s. An internal explosion occurs, splitting the object into two equal chunks and adding 16 J of kinetic energy to the system. What is the average acceleration of the two chunks if the explosive separation takes place over a 0.16-s time interval? ●●
37. A railroad car that has an inertia of 5.04×10^4 kg is moving to the right at 4.25 m/s when it collides and couples with three identical railroad cars that are already coupled together and moving in the same direction at 2.09 m/s. If the acceleration of the single car is 5.45 m/s^2 to the left while it is changing velocity, how long does the interaction take? ●●
38. A 90-kg halfback on a football team runs head-on into a 120-kg opponent at an instant when neither has his feet on the ground. The halfback is initially going west at 10 m/s, and his opponent is initially going east at 4.37 m/s. The collision is totally inelastic. (*a*) If the collision takes 0.207 s, what is the acceleration of each player? (*b*) How much kinetic energy is converted to incoherent energy during the collision? ●●
39. You are watching a hockey game on your digital video recorder. Your team's goalie is at rest when he catches a 0.16-kg puck moving straight toward him. The announcer says it was a 40-m/s slap shot, based on radar gun data. Normally the goalie would be braced, but in this case he is resting on parallel skate blades so that there is negligible friction to impede his subsequent motion. The collision is too fast to observe quantitatively, but by running the video in slow motion, you can tell that the goalie's glove is 150 mm from his body when

the puck strikes and is right against his chest after the successful catch. It is also clear that the stunned goalie (holding the puck) drifts backward at 71 mm/s for a second or two after the collision is over. (*a*) Compute the average acceleration of the puck and of the goalie during the collision. (*b*) Determine the inertia of the goalie (including his equipment). (*c*) How much kinetic energy is converted to thermal energy in the glove? Check your answers by solving in two ways when possible. •••

40. Two objects, one red and one yellow, start from rest and interact with each other along an x axis. (*a*) If $m_{red} < m_{yellow}$, what are the ratios of their changes in momentum, velocity, and kinetic energy? (*b*) Compare these ratios with the ratio of their accelerations. (*c*) What does your answer to part *a* tell you about the relative changes of these quantities for a ball and Earth when you drop the ball? •••

7.8 Nondissipative interactions

41. Draw a bar diagram representing the various kinds of energy for the cart-spring system in Figure 7.26*b* when the end of the cart attached to the spring is at locations x_1, x_2, and x_3. •

42. The potential energy of an interaction is given by $U(x) = ax^2$, where $a = +6.4\ \text{J/m}^2$. (*a*) If the initial speed of a 0.82-kg object in this system is 2.23 m/s at $x = 0$, how far does the object travel before it reaches a speed of $v = 0$? (*b*) Does your answer in part *a* depend on whether the object is traveling in the positive or negative x direction? •

43. Suppose we have a block mounted on a spring and define the zero point for the elastic potential energy of the system to be the relaxed position (meaning the block is positioned such that the spring is neither compressed nor stretched). What is the appropriate sign of the elastic potential energy when the block is repositioned so that the spring is (*a*) stretched and (*b*) compressed? (Hint: Think about what happens to the elastic potential energy when the stretched or compressed spring is released.) ••

44. A 0.36-kg cart and a 0.12-kg cart are held together with a compressed spring between them. When they are released, the 0.36-kg cart moves at 1.1 m/s to the right. How much elastic potential energy was stored in the spring before the release? ••

45. A 0.530-kg cart moving at 0.922 m/s to the right collides elastically with a 0.25-kg cart initially at rest. The 0.25-kg cart then moves off rapidly and compresses a spring before the 0.530-kg cart can catch it again. At the instant the 0.25-kg cart comes to rest, what is the energy stored in the spring? ••

46. A low-friction track is set up with a spring on the right end. In the middle of the track, a 2.0-kg cart moving to the right at 5.0 m/s overtakes a 3.0-kg cart moving to the right at 2.0 m/s. After the collision, the 2.0-kg cart is moving at 1.7 m/s. (*a*) What is the coefficient of restitution of the collision? (*b*) How much energy is stored in the spring when it is at its maximum compression after the collision? ••

47. A 5-kg object is subject to an interaction that has a potential energy $U(x) = \frac{1}{2}kx^2 - bx$, where $k = 2.0\ \text{J/m}^2$ and $b = 1.5\ \text{J/m}$. At $x = 1.0$ m, the object is found to be moving to the left at 4 m/s. What is the speed of the object at $x = 2.0$ m? What is the object's kinetic energy at $x = 3.0$ m? ••

48. You release a block from the top of a long, slippery inclined plane of length ℓ that makes an angle θ with the horizontal. The magnitude of the block's acceleration is $g \sin\theta$ (see Section 3.7). (*a*) For an x axis pointing down the incline and having its origin at the release position, derive an expression for the potential energy of the block-Earth system as a function of x. (*b*) How would your expression change if you put the origin of the x axis at the bottom of the incline? (*c*) Does the location of the origin make any difference in the dynamics of the situation? (*d*) Use the expression you derived in part *a* to determine the speed of the block at the bottom of the incline. Does your speed value agree with the result obtained by a kinematic calculation? •••

49. On a low-friction track, a 0.36-kg cart initially moving to the right at 2.05 m/s collides elastically with a 0.12-kg cart initially moving to the left at 0.13 m/s. The 0.12-kg cart bounces off the 0.36-kg cart and then compresses a spring attached to the right end of the track. (*a*) At the instant of maximum compression of the spring, how much elastic potential energy is stored in the spring? (*b*) If the spring then returns all of this energy to the cart, and the two carts again collide, what is the final velocity of each cart? •••

7.9 Potential energy near Earth's surface

50. A rock climber accidentally drops a 4.5-kg backpack, and it falls 160 m to the ground below. What is the change in the gravitational potential energy of the system comprising the backpack and the Earth? •

51. At graduation, you toss your mortarboard cap straight up at some initial speed v. How fast is it moving when it comes back down to your hands if you (*a*) ignore air resistance and (*b*) include it? How do you know? •

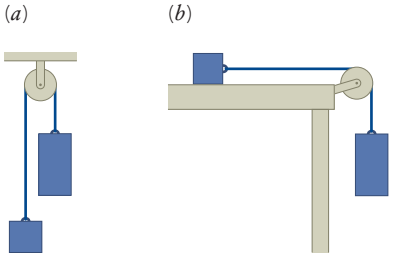
52. An office worker has his lunch on the 53rd floor of an office building. In which case is the gravitational potential energy of the lunch-Earth system greatest: (*a*) the worker carried the lunch up to the 53rd floor in the elevator, (*b*) he carried it up the stairs to the 53rd floor, or (*c*) he got a ride in a helicopter to the top of the 68-floor building and then carried the lunch down in the elevator to the 53rd floor? •

53. The fastest baseball pitchers can throw the 0.145-kg ball at speeds of about 45 m/s. Ignoring air resistance, what height must the ball be dropped from to hit the ground at this speed? •

54. A typical loaded commercial jet airplane has an inertia of 2.1×10^5 kg. (*a*) Based on the time intervals involved, which do you expect to require more energy: getting the plane up to cruising speed or getting it up to cruising altitude? (*b*) How much energy does it take to get the plane to a cruising speed of 270 m/s? (Ignore wind drag.) (*c*) How much energy does it take to get the plane to a cruising altitude of 10,400 m if it travels to that altitude at its (constant) cruising speed? ••

55. Two 2.4-kg blocks are connected by a string draped over the edge of a slippery table, so that one block is on the table and the other is just hanging off the edge. A restraint holds the block on the table in place, and the string is 0.50 m long. After the restraint is removed, what speed does the pair of blocks have at the instant the upper one is pulled off the table? ●●
56. A 30-kg child jumps to the ground from the top of a fence that is 2.0 m high. You analyze the problem using upward as the positive x direction. (*a*) Taking $x = 0$ to be at the bottom of the fence, what are the initial potential energy of the child-Earth system and the change in the system kinetic energy during the jump? (*b*) Repeat part *a* for $x = 0$ at the top of the fence. ●●
57. You lean out of your dorm window, which is 12 m above the ground, and toss a 0.12-kg ball up to a friend at a window 11 m above you. (*a*) What is the slowest initial speed at which you can throw the ball so that he just catches it? (*b*) If you throw the ball at this speed and your friend misses it, what is its kinetic energy the instant before it hits the ground? (*c*) Suppose you throw the ball directly downward at the speed calculated in part *a*. What is its kinetic energy just before it hits the ground? ●●
58. Two blocks are hung by a string draped over a pulley, a 1.0-kg block on the left and a 3.0-kg block on the right. The two blocks start out at rest and at the same height. (*a*) What is the change in the gravitational potential energy of the system of blocks and Earth when the 3.0-kg block has dropped 0.53 m? (*b*) What is the change in the kinetic energy of the system between release and this instant? (*c*) What is the velocity of the 1.0-kg block at this instant? ●●
59. You drop a rubber ball from a height of 3.0 m. It bounces off a concrete surface to a height of 2.7 m. (*a*) What is the coefficient of restitution for this collision? (*b*) You want to get the ball to bounce upward to a height of 7.3 m. From the same starting point, how fast must you throw the ball, and in which direction? ●●
60. A 100-g apple is dropped from a height of 12 m and 1 second later is struck by a 100-g arrow flying upward at 15 m/s. (*a*) What is the speed of the apple and the arrow just after collision? (*b*) How much energy is lost in the collision? (*c*) How long does it take the combined object to reach the ground from the time the apple was dropped? ●●
61. A uniform chain of inertia m and length ℓ is lying on a slippery table. When one quarter of its length hangs over the edge, the chain begins to slip off. How fast is it moving when the last bit of it slips off the table? (Ignore friction.) ●●●
62. A uniform chain of inertia m and length ℓ is lying on a slippery table. When just the tip hangs over the edge, the chain begins to slip off. (Ignore friction.) Calculate the speed of the chain as a function of time. (Hint: $\int \frac{dx}{v} = \int dt$. What's the connection between x and v?) ●●●
63. Two blocks are connected by a string draped over a small pulley. The smaller block has half the inertia of the larger block. For which configuration, (*a*) or (*b*), in Figure P7.63 are the speeds of the blocks greatest after the larger block has descended a distance d? Ignore friction between the table and the sliding block, and check your answer with a computation. ●●●

Figure P7.63



7.10 Dissipative interactions

64. Gravity-powered roller coasters have a motorized chain assembly that hauls the cars up to the top of the first hill. No additional source energy is supplied for the rest of the trip. What is the maximum possible height for all subsequent hills? ●
65. While you are standing on your balcony 8 m above the ground, your friend tosses a 0.4-kg book at you from the ground at 14 m/s. The book barely makes it to your hand. How much energy was dissipated by air resistance? ●
66. A 5.3-kg steel ball is dropped from a height of 20 m into a box of sand and sinks 0.20 m into the sand before stopping. How much energy is dissipated through the interaction with the sand? ●●
67. Two cars collide inelastically on a city street. For the two-car system, which of the following are the same in any inertial reference frame: (*a*) the kinetic energy, (*b*) the momentum, (*c*) the amount of energy dissipated, (*d*) the momentum exchanged? ●●
68. A 0.70-kg basketball dropped on a hardwood floor rises back up to 65% of its original height. (*a*) If the basketball is dropped from a height of 1.5 m, how much energy is dissipated in the first bounce? (*b*) How much energy is dissipated in the fourth bounce? (*c*) To which type of incoherent energy is the dissipated energy converted? ●●
69. In a simple throwball game on ice, a 90-kg athlete throws a 1.0-kg ball at 15 m/s. The ball is caught by his 80-kg teammate. If 20% of the source energy is lost in the initial throw, then what are the recoil speeds of the athletes? What is the source energy? ●●
70. A 2.2-kg measuring instrument is mounted on a balloon by your scientific team for atmospheric studies. At the top of its flight, the instrument is released from the balloon and falls most of the way back to Earth before a parachute opens. You are told that the magnitude of the acceleration at any instant t before the parachute opens is given by $|a| = ge^{-t/\tau}$, where g is the acceleration due to gravity, e is the base of natural logarithms, and τ is a time constant that depends on the shape of the instrument and in this case is 5.68 s. Your primary concern is how much the instrument heats up as it falls, due to air

resistance. At what rate, in joules per second, is energy dissipated before the parachute opens? Express your answer as a function of t, where $t = 0$ at release. (Hint: Integrate the acceleration to calculate speed and displacement.) •••

71. A nursery rhyme goes "Humpty Dumpty sat on a wall, Humpty Dumpty had a great fall. All the king's horses and all the king's men couldn't put Humpty Dumpty together again." (*a*) Describe the energy of the Earth-HD system before, during (including the instant just before impact), and after the fall. Use bar diagrams showing K, U, E_s, and E_{th} to illustrate your description. (*b*) In terms of the energy of the Earth-HD system, explain why all the king's horses and all the king's men could not put HD together again. (*c*) Now assume that things had turned out differently for HD. The king's men were able to set up a net so that he fell into the net and bounced back up and once again sat on the wall. Draw energy bar diagrams for this situation, showing the distribution of energy during the fall, at the instant HD arrives at the net (choose the instant at which the net with HD in it is stretched to the lowest position), during the ascent, and when he is back sitting on the wall. •••

72. A small block of wood of inertia m_{block} is released from rest a distance h above the ground, directly above your head. You decide to shoot it with your pellet gun, which fires a pellet of inertia m_{pellet}. After the block has fallen a distance d, the pellet hits it and becomes embedded in it, kicking it upward. At the instant of impact, the pellet is moving at speed v_{pellet}. (*a*) To what maximum height h_{max} does the pellet-block system rise above the ground? Express your answer in terms of the given variables and known values, such as g. (*b*) How much energy is dissipated in the collision? •••

Additional Problems

73. Is the collision between two billiard balls a nondissipative interaction? (Hint: What does the fact that you can hear the collision tell you?) •

74. Draw before and after energy bars for the collision shown in Figure 6.8*a* and 6.8*b*. •

75. A 60-g Mars bar will supply you with 1095 kJ of energy. How many of them would you need to get enough energy to climb the first 800 m of the highest structure in the world? The next 800 m? Assume your inertia to be 80 kg. •

76. You drop a ball from some height. At which position do you expect its speed to be one-half the speed at which it eventually hits the ground: higher than the midpoint of the path, at the midpoint, or lower than the midpoint? Does air resistance change your answer? ••

77. Is it impossible to convert incoherent energy to coherent energy? Can you think of an example where thermal energy can (at least partially) be converted to coherent energy? ••

78. People unenlightened about physics sometimes fire a gun straight up into the air to celebrate big events. By discussing the forms of energy at each stage of the process, explain why this is a bad idea. ••

79. A 1-kg cart and a 2-kg cart are held together with a coupler that contains a small charge. The charge is exploded and sends the 1-kg cart rolling away at +4.0 m/s. What is the speed of the 2-kg cart? How much energy does the explosion create if 70% of it is lost as heat? ••

80. Wind turns a windmill that stores energy in a battery. Assume that the windmill turns at a constant rate while the battery is charging. The battery is then used to run a toy car that moves at a constant speed until the battery goes dead. Draw bar diagrams representing a sequence of instants that depict the various conversions of energy from one form to another in this situation. For each conversion, name the system you are describing. ••

81. You and your bike collide at 12 m/s into the back of a car stopped at a red light. The car does not have its brakes applied and so is jolted forward. The driver immediately leaps out crying, "Whiplash!" Facing a day in court, you have to determine the acceleration of the car as a result of the collision. You note that you and your bike (combined inertia 80 kg) came to a complete halt in the collision and that the rim of your front wheel was pushed all the way to the center hub—the same distance the car traveled during the crash. The diameter of the bike wheel before the crash was 0.75 m, and the car's inertia is 1800 kg. ••• CR

82. An amusement park commissions your company to design a roller coaster. Your company specializes in an approach in which the cars are accelerated on a horizontal section of track by magnetic motors rather than being hauled up to the top of a hill and released. The park officials would like the riders to be accelerated via your horizontal design and then sent straight up a vertical length of track 66.4 m tall. (After the cars pass this high point, the ride is the regular ups and downs of a conventional roller coaster.) Your calculations tell you that 212 motors will get the loaded cars up to the necessary speed at an acceleration of $0.85g$, but you wonder whether or not the park has enough horizontal distance available for the acceleration device. ••• CR

83. A 1.00-kg cart has attached to its front end a device that explodes when it hits anything, releasing a quantity of energy E. This cart is moving to the right with speed v when it collides head-on with a 2.00-kg cart traveling to the left at the same speed v. The explosive goes off when the carts hit, causing them to rebound from each other. If one-quarter of the explosive energy is dissipated into the incoherent energy of noise and deformation of the carts, what is the final velocity of each cart? Solve in the zero-momentum reference frame and then check your result by solving in the Earth reference frame. •••



8

Force

CONCEPTS

8.1 Momentum and force
8.2 The reciprocity of forces
8.3 Identifying forces
8.4 Translational equilibrium
8.5 Free-body diagrams
8.6 Springs and tension

QUANTITATIVE TOOLS

8.7 Equation of motion
8.8 Force of gravity
8.9 Hooke's law
8.10 Impulse
8.11 Systems of two interacting objects
8.12 Systems of many interacting objects

At this point in your study of physics, you can already solve a wide range of problems using just the conservation laws. Up to now we have applied these laws in only isolated or closed systems, however, and it is not always possible to identify such a system when working on problems. For this reason, you also need to know an approach to problem solving that is valid in systems that are interacting with their surroundings. If you look back at the material in earlier chapters, you will notice that many of the interactions we discussed involve pushing, pulling, or rubbing—all actions related to the everyday notion of *force*. So, in order to be able to deal with systems that are not isolated, you must know about the concept of force.

The term *force* is a familiar one—usually associated with the capacity to change the motion, or non-motion, of objects (for example, pushing a chair along the floor) or to cause physical change (say, crushing an empty can). To which of the physical quantities I have defined so far does this notion of force best correspond, however, and how can we quantify it?

8.1 Momentum and force

To relate the intuitive concept of force to the quantities we have encountered so far, consider the following scenario. You are on a bicycle coming down a steep hill and approaching a sharp turn at 30 km/h. Suddenly you discover that your brakes are not working, leaving you no option but to slam into a concrete wall. The force of impact will be considerable, and if you were going faster—say, 40 km/h—it would be even greater. Conversely, if you were going at a mere 3 km/h, the impact might not be that bad. And it is not just your speed that matters. When you are carrying a backpack full of bricks and traveling at 30 km/h, the force of impact is greater than when you are traveling at the same speed but carrying nothing on your back.

Realizing that both your speed and your inertia govern the force of the impact you make with the wall, you may be tempted to believe that force is related solely to change in momentum, but things are not that simple. To see why, imagine that part of the concrete wall is padded with mattresses. Which part do you head for? You don't need any physics to realize that the mattresses offer you protection. When you hit a mattress, the force of impact is smaller than when you hit the concrete. In terms of momentum change and energy change, however, it doesn't make any difference which one you choose. In either case, your speed decreases by the same amount, 30 km/h, and so the decreases in your momentum and kinetic energy are the same no matter how you come to a stop.

The difference between hitting a mattress and hitting the concrete is that the mattress changes your momentum over a longer time interval. The momentum change therefore occurs at a slower rate. Evidently, the longer the interaction time interval, the smaller the force of impact. In other words, what determines the magnitude of the force of impact is not the absolute value of your momentum change but rather the rate at which this change happens. The faster your momentum changes, the larger the force of the impact on you.

This scenario illustrates two important points about forces. First, forces are manifestations of interactions—the force exerted by the mattress on you, for example, is one part of the interaction between you and the mattress (another part is the force exerted by you on the mattress). Second, for an object that is participating in one interaction only, we can quantitatively define **force:**

> **The force exerted on the object is the time rate of change in the object's momentum.**

Because it is related to a change in momentum, force is a vector and so has direction as well as magnitude. The direction of any given force is the same as the direction of the momentum change the force causes. When you hit the wall in our bicycling example, for instance, the direction of the force exerted by the wall on you is opposite your direction of travel because the direction of your momentum change is opposite your direction of travel.

Even though the arguments leading up to the relationship between force and time rate of change in momentum all sound very plausible, I should warn you that the consequences of this relationship differ from your intuitive notions about forces in a number of ways.

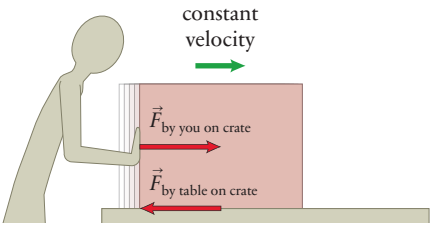
CONCEPTS



8.1 Imagine pushing a crate in a straight line along a surface at a steady speed of 1 m/s. What is the time rate of change in the momentum of the crate?

The answer to this checkpoint should surprise you. After all, you *are* pushing the crate. So how can the time rate of change in the momentum of the crate be zero? The problem here is that the crate is interacting not just with you but also with the surface because there is friction between it and the surface. The directions and magnitudes of these two interactions (between the crate and your hands and between the crate and the surface) are such that their combined action causes no change in the momentum of the crate, and so it continues to move at constant speed in the direction of your push. In each interaction, there is a force exerted on the crate—exerted by you in one case and by the surface in the other case. The force exerted by your hands on the crate tends to move the crate

Figure 8.1 When you push a crate along a surface, two opposing forces are exerted on the crate: one by you (the push), the other by the surface (friction). If these forces are equal in magnitude, the crate moves at constant velocity.



forward, while the force exerted in the horizontal direction by the surface on the crate (the *force of friction* or *frictional force*) tends to resist the forward motion of the crate. If the two opposing forces exerted on the crate are equal in magnitude, their vector sum (sometimes called the *net force*) is zero. In other words, if the surface pushes on the crate just as hard as you push on it in the opposite direction, the acceleration effects of the two opposing forces cancel and the crate moves at constant velocity (**Figure 8.1**). If the force exerted by the surface on the crate is smaller than the force you exert on it, the crate speeds up; if the force exerted by the surface is larger, the crate slows down.

For an object A interacting with more than one other object (such as a crate interacting both with you and with the surface), the resulting effect on A's momentum is due to the combined action of a number of forces, one for each interaction. Experimentally, we discover that forces obey the *principle of superposition*: The vector sum of the forces exerted on A is equal to the vector sum of the rate of change caused in A's momentum by the individual forces. Consequently:

The vector sum of all forces exerted on an object equals the time rate of change in the momentum of the object.

CONCEPTS

For an object of fixed inertia, a change in momentum implies a change in velocity. So, when the vector sum of the forces exerted on an object is not zero, the object has a nonzero acceleration. The vector sum of the forces exerted on an object points in the same direction as the object's acceleration.

8.2 Imagine pushing on a crate initially at rest so that it begins to move along a floor. (*a*) While you are setting the crate in motion, increasing its speed in the desired direction of travel, what is the direction of the vector sum of the forces exerted on it? (*b*) Suppose you suddenly stop pushing and the crate slows to a stop. While the crate slows down, what is the direction of the vector sum of the forces exerted on it? (*c*) What is the direction of the vector sum of the forces once the crate comes to rest?

8.2 The reciprocity of forces

Another surprising aspect of the definition of force is that, because of the reciprocal nature of interactions, forces always come in pairs: When two objects interact, each exerts a force on the other. To see this, consider **Figure 8.2**, which illustrates two elastic collisions. In both collisions, cart 1 of inertia $m_1 = 0.12\ \text{kg}$ is initially at rest and cart 2 of inertia $m_2 = 0.24\ \text{kg}$ is initially moving at 0.60 m/s. For both collisions, the final velocity of cart 1 is 0.80 m/s and that of cart 2 is 0.20 m/s. The only difference between the two collisions is that in Figure 8.2*a* the impact is softened by a spring, much like the cushioning of a fall by a mattress. As we did before, during the collision, we can ignore all interactions except that between the two carts.

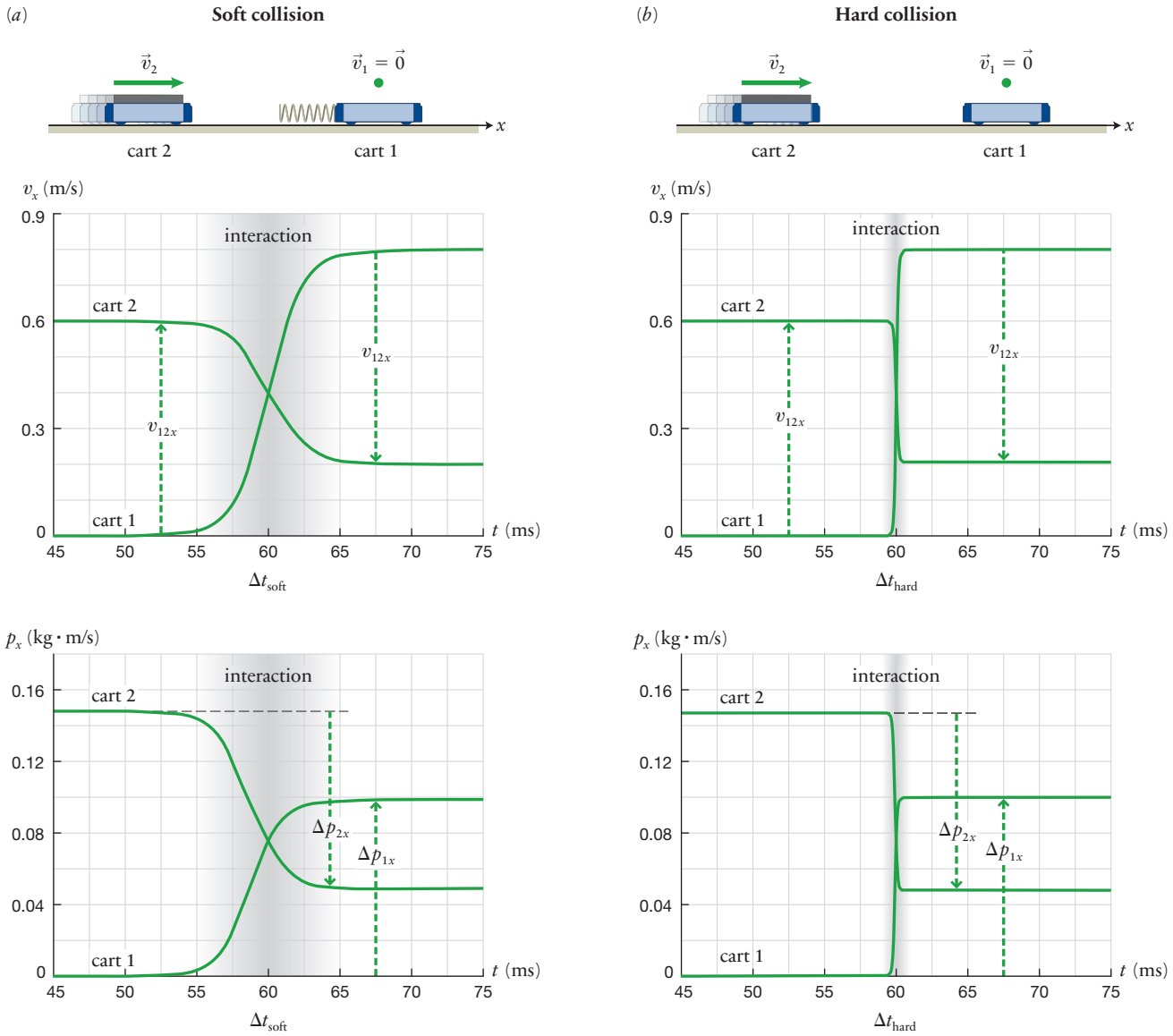
8.3 (*a*) Verify that for both collisions in Figure 8.2 the momentum of the two-cart system remains constant. (*b*) Verify that both collisions are elastic.

As this checkpoint shows, in both collisions the momentum change of cart 1 is compensated for by an opposite change in the momentum of cart 2. Let's now relate these momentum changes to forces. For the soft collision, the changes in momentum take place during a time interval Δt_{soft} (about 10 ms). During this time interval, the x component of the momentum of cart 1 changes by an amount $\Delta p_{1x} = +0.096\ \text{kg}\cdot\text{m/s}$. Because force equals time rate of change in momentum, the x component of the average force exerted by cart 2 on cart 1 during the time interval Δt_{soft} is $(F_{\text{by 2 on 1}})_x = \Delta p_{1x}/\Delta t_{\text{soft}} = (+0.096\ \text{kg}\cdot\text{m/s})/(0.010\ \text{s}) = +9.6\ \text{kg}\cdot\text{m/s}^2$.

Because the two-cart system is isolated, the change in the momentum of cart 2 must be equal in magnitude to that of cart 1 and in the opposite direction, and so $\Delta p_{2x} = -0.096\ \text{kg}\cdot\text{m/s}$. Because the time interval it takes for this change to occur is also Δt_{soft}, we calculate for the x component of the time rate of change in the momentum of cart 2, $(F_{\text{by 1 on 2}})_x = \Delta p_{2x}/\Delta t_{\text{soft}} = -9.6\ \text{kg}\cdot\text{m/s}^2$. In words, the force exerted by cart 1 on cart 2 is equal in magnitude to the force exerted by cart 2 on cart 1 and in the opposite direction.

For the hard collision in Figure 8.2*b*, the momentum changes are the same as for the soft collision, but now the time interval during which these changes take place is much shorter, $\Delta t_{\text{hard}} \approx 1.0$ ms. As you might expect, the forces are much larger: For cart 1, we now have $(F_{\text{by 2 on 1}})_x = \Delta p_{1x}/\Delta t_{\text{hard}} = (+0.096\ \text{kg}\cdot\text{m/s})/(0.0010\ \text{s}) = +96\ \text{kg}\cdot\text{m/s}^2$. The force exerted by cart 2 on cart 1 is again equal in magnitude to the force exerted by cart 1 on cart 2 and in the opposite direction because the two-cart system is isolated: $\Delta p_{1x} = -\Delta p_{2x}$. Now $(F_{\text{by 1 on 2}})_x = \Delta p_{2x}/\Delta t_{\text{hard}} = -96\ \text{kg}\cdot\text{m/s}^2$. So we conclude:

Figure 8.2 (*a*) Soft and (*b*) hard collisions between two carts on a low-friction track. The inertias are $m_1 = 0.12$ kg and $m_2 = 0.24$ kg.



Whenever two objects interact, they exert on each other forces that are equal in magnitude but opposite in direction.

The pair of forces that two interacting objects exert on each other is called an **interaction pair.**

8.4 Does the conclusion just stated apply to inelastic collisions?

8.5 If you drop a book from a certain height, it falls (accelerating all the while) because of the gravitational force exerted by Earth on it. Because forces always come in interaction pairs, the book must exert a force on Earth. (*a*) How does the magnitude of $F_{\text{by book on Earth}}$ compare with the magnitude of $F_{\text{by Earth on book}}$? (*b*) Is Earth accelerating during the interaction? Why or why not?

The conclusion that interacting objects exert equal forces in opposite directions on each other is a direct result of the law of conservation of momentum and of our definition of force. This reciprocity of forces means, for instance, that if you push a crate along the floor, the crate pushes back on you with a force of equal magnitude. Can an inanimate object like a crate exert a force, and, if so, how do we know this force exists?

Let's settle the first question first, using the concrete wall of our earlier example. Even though the wall is inanimate, there is no denying it exerts a force on you when you slam into it. I don't recommend trying it, but you can imagine feeling this force. So inanimate objects can exert forces.

Now for the second question: How do we know the crate exerts a force on you when you push it? To see

CONCEPTS

that objects push back when you push on them, imagine leaning against a door when someone suddenly opens it away from you—you tumble into the door opening precisely because the force exerted by the door on you (a force that was keeping you upright) is suddenly removed. The force exerted by you on the door and the force exerted by the door on you form an *interaction pair*.

Here is another example: Imagine tipping over a chair by kicking it with your bare foot. The chair is tipped over by the force exerted by your foot on it. Your foot hurts because of the other half of the interaction pair—the force exerted by the chair on your foot.

EXAMPLE 8.1 Splat!

A mosquito splatters onto the windshield of a moving bus and sticks to the glass. (*a*) During the collision, is the magnitude of the change in the mosquito's momentum smaller than, larger than, or equal to the magnitude of the change in the bus's momentum? (*b*) During the collision, is the magnitude of the average force exerted by the bus on the mosquito smaller than, larger than, or equal to the magnitude of the average force exerted by the mosquito on the bus?

❶ **GETTING STARTED** Because the mosquito sticks to the windshield, the bus and mosquito move together after impact and so this event is a totally inelastic collision. I therefore choose the mosquito plus the bus as my system.

❷ **DEVISE PLAN** During the collision, the system is isolated, and so its momentum doesn't change. I can use this fact to compare the magnitudes of the changes in momentum. The average forces exerted on the mosquito and the bus are each equal to the corresponding momentum change divided by the time interval over which that momentum change takes place. I also know that this time interval is the same for the mosquito and the bus.

❸ **EXECUTE PLAN** (*a*) Because the system's momentum is unchanged, I know that $\Delta\vec{p}_{\text{mos}} = -\Delta\vec{p}_{\text{bus}}$ and so $|\Delta\vec{p}_{\text{mos}}| = |\Delta\vec{p}_{\text{bus}}|$. ✔

(*b*) Because the changes in momentum are equal in magnitude and because the time interval over which the change takes place is the same for both objects, the magnitudes of the average forces must also be equal: $|\vec{F}_{\text{by mos on bus, av}}| = |\vec{F}_{\text{by bus on mos, av}}|$. ✔

❹ **EVALUATE RESULT** Although the answer to part *a* conforms to what I've seen before, the answer to part *b* is surprising. Intuitively I expect the magnitude of the average force exerted by the bus to be much larger than the magnitude of the average force exerted by the mosquito because the mosquito is crushed while nothing happens to the bus. However, I know that the force magnitude required to crush the mosquito is very small, and such a small force magnitude would have no effect on the bus. It therefore makes sense that the two forces are of equal magnitude.

CONCEPTS

8.6 In Example 8.1, the mosquito has an inertia of 0.1 g and is initially at rest, while the bus, with an inertia of 10,000 kg, has an initial speed of 25 m/s. The collision lasts 5 ms. (*a*) Calculate the final speeds of the mosquito and the bus. (*b*) What is the average acceleration of each during the collision? (*c*) By how much does the momentum of each change? (*d*) During the collision, how large is the average force exerted by the bus on the mosquito? (*e*) How large is the average force exerted by the mosquito on the bus?

8.3 Identifying forces

To identify the forces exerted on an object, it is convenient to distinguish between contact forces and field forces:

- **Contact forces** are forces that arise when objects physically touch each other. This category of forces includes forces due to pushing, pulling, and rubbing.
- **Field forces** are forces associated with what is called "action at a distance." In this case, the objects exerting forces on each other do not need to be physically touching. For any object larger than atoms, gravitational and electromagnetic forces are the only field forces.

Identifying contact forces is easy: Whenever the surfaces of two objects touch, the objects exert contact forces on each other. Nearly all familiar forces fall into this category: the pull of a rope, the traction of a tire on a road, the force that keeps an airplane in the air, the push of a bumper, the hit of a hammer, the force that keeps you afloat when you are swimming, the drag caused by air resistance, and so on.

To determine which field forces to include in a given situation, gravitational and electromagnetic interactions must be considered. Gravitational interactions take place between any two objects. In our daily life, for instance, we each interact with Earth via the gravitational force, as does any other object—this book,* an apple, a house, and so on. The gravitational interaction we have with one another and with all other objects except Earth, however, is negligibly small. If you hold this book with your arms stretched out in front of you and let it go, it accelerates toward Earth, not toward you.

Because ordinary objects tend to be both electrically neutral and nonmagnetic, we usually do not need to

*Hold a book with your arms stretched out in front of you and move it up and down. Do you feel the gravitational force exerted by Earth on the book? You can visualize this interaction as an invisible stretched spring reaching from the center of Earth to the book. Imagine stretching and relaxing this "gravitational spring" as you move the book up and down.

consider field forces due to magnetic and electric interactions, unless a problem specifically mentions these types of interactions. We will treat electric and magnetic forces in detail in later chapters.

So, in most cases the only field force we need to consider is the gravitational force exerted by Earth. Only if a problem specifically mentions magnetic or electric interactions must we also take the field forces due to these interactions into account.

EXERCISE 8.2 What forces?

Identify all the forces exerted on the italicized object in each situation: (*a*) A *book* is lying on top of a magazine on a table. (*b*) A *ball* moves along a trajectory through the air. (*c*) A person is sitting on a *chair* on the floor of a room. (*d*) A *magnet* floats above another magnet that is lying on a table.

SOLUTION Because the interaction with the air is usually very small, I'm ignoring this interaction in this problem.

(*a*) Contact forces: The book is in contact with the magazine only, and so the magazine exerts a contact force on the book. Field forces: Earth exerts a gravitational force on the book. ✔

(*b*) Contact forces: The ball is not in contact with anything during its flight, and so there are no contact forces. Field forces: There is a gravitational interaction between Earth and the ball, so Earth exerts a gravitational force on the ball. ✔

(*c*) Contact forces: The chair is in contact with the floor and the person, and so both the floor and the person exert a force on the chair. Field forces: Earth exerts a gravitational force on the chair. ✔

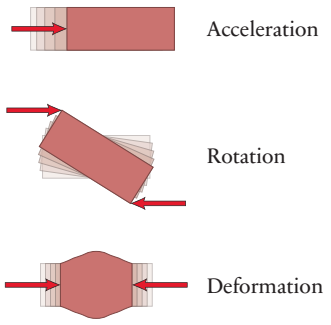
(*d*) Contact forces: The magnet floats, so it is not in contact with anything and thus there are no contact forces. Field forces: Both gravitational and magnetic forces occur in this problem: a gravitational force exerted by Earth on the magnet and a magnetic force exerted by the second magnet. ✔

8.7 A magnet lies on a table. You place a second magnet near the first one so that the two repel each other. Identify all the forces exerted on the first magnet.

8.4 Translational equilibrium

An object or system whose motion or state is not changing is said to be in **equilibrium.** For example, an object at rest or moving at constant velocity—in other words, its momentum is constant and its acceleration is zero—is said to be in **translational equilibrium.** According to the definition of force given in Section 8.1, constant momentum implies that the vector sum of the forces exerted on the object must be zero. Note that being in translational equilibrium doesn't necessarily mean that no forces are exerted on the object. It means only that all the force vectors add up to zero and therefore do not change the object's momentum.

Figure 8.3 Forces can have various effects on an object.



As **Figure 8.3** illustrates, however, forces can also cause an object to rotate or to deform. Which combination of these effects occurs depends on where and how forces are exerted on the object. For now we consider only the acceleration caused by forces. Deformation is discussed in Chapters 7 and 8 and rotation in Chapter 11.

A book at rest on the floor is in translational equilibrium. So, too, is the crate pushed at constant velocity in Figure 8.1. In both of these cases, as with a coil spring squeezed between your hands, forces are exerted even though the object is in translational equilibrium.

Whenever an object is at rest or moving at constant velocity, the vector sum of the forces exerted on the object is zero, and the object is said to be in translational equilibrium.

As an example of an object in translational equilibrium, consider a book lying on a table, as in **Figure 8.4*a***. The book's momentum isn't changing, and therefore the vector sum of the forces exerted on it must be zero. What forces are exerted on the book? The only field force we need to consider is gravity. Even though the book is at rest, we know it is subject to gravity because if we suddenly remove the table from under it, the book accelerates downward (Figure 8.4*b*). Because the book is touching the table, we know there is a contact force exerted by the table on the book—the force that holds the book up and prevents it from falling through the table.

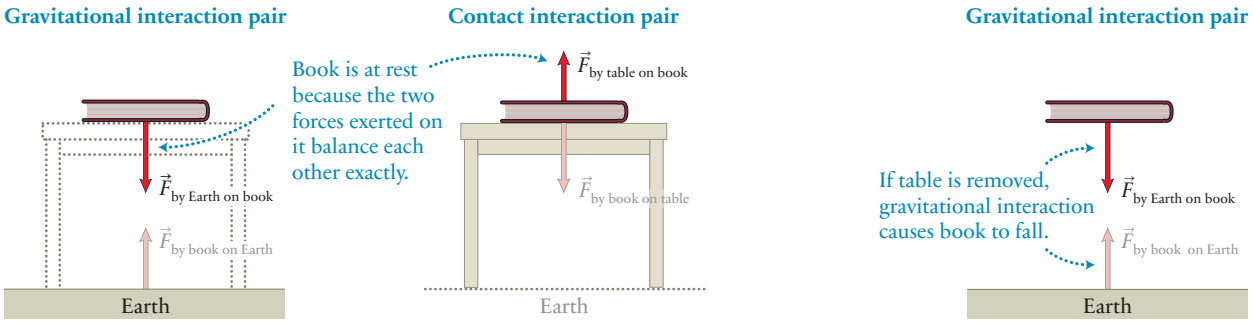
These are the only two forces exerted on the book. Because the vector sum of these forces must be zero, we conclude that the contact force exerted by the table on the book must be of the same magnitude as the gravitational force exerted by Earth on the book but in the opposite direction.

An object is also in translational equilibrium when it is moving at constant velocity. An example is the crate of Figure 8.1. Constant velocity means no change in momentum, and therefore the vector sum of the forces exerted on the crate must be zero. Considering only forces in the direction of motion, we determine two contact forces exerted on the crate: one by you pushing and one by the floor rubbing. In order for the

Figure 8.4

(*a*) Book at rest on table (vector sum of forces exerted on book is zero)

(*b*) Book in free fall (vector sum of forces exerted on book is not zero)



vector sum of these forces to be zero, the frictional force exerted by the floor on the crate must be of the same magnitude as the force exerted by you on the crate but in the opposite direction.

8.8 In Figure 8.4, are the contact force exerted by the table on the book and the gravitational force exerted by Earth on the book an interaction pair?

8.5 Free-body diagrams

Knowing about forces gives you a powerful tool for analyzing physical situations:

> **If the vector sum of the forces exerted on an object is known, the time rate of change in the object's momentum is also known, and so the object's subsequent motion can be calculated.**

Note the condition: To calculate the object's motion you need to know the vector sum of the forces exerted **on** the object only. The forces exerted **by** the object on its surroundings do not enter into the picture because they do not contribute to the change in the object's momentum. So, if you are considering an accelerating car and you need to calculate the time rate of change of its momentum, then you need to determine the vector sum of the forces exerted *on* the car (by the road, by the air, by gravity, and so on). The forces exerted *by* the car (on the road, on the air, . . .) determine what happens to those objects and not what happens to the car.

Whenever you are interested in one object in a collection of interacting objects, you must separate that one object and the forces exerted on it from the rest of the collection of objects and ignore the forces exerted by that object on anything else. To facilitate this separation, you should use what is called a **free-body diagram.** The procedure for drawing such diagrams is explained in "Drawing free-body diagrams." It is often also helpful to make a small sketch to identify the objects of interest. Be sure, however, to keep this "situation sketch" and your free-body diagram separate.

Until we begin to treat motion in more than one dimension in Chapter 10, we shall consider only the forces that are exerted in the direction of motion. For example, when pushing a cart on a horizontal track, you need only draw the forces along the horizontal direction. The forces along the vertical direction—an upward contact force exerted by the track on the cart and a downward gravitational force exerted by Earth on the cart—add up to zero and cause no change in the cart's motion.

The importance of free-body diagrams can hardly be exaggerated: Without a good free-body diagram, it is very difficult to analyze correctly any problem involving forces. In Section 4.4, I pointed out that separating a system from its environment is the most fundamental step in analyzing physical problems. Free-body diagrams help make this separation explicit. Let's therefore put the procedure into practice in a few examples.

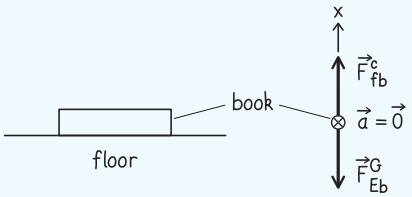
CONCEPTS

EXERCISE 8.3 A book on the floor

Draw a free-body diagram for a book lying motionless on the floor.

SOLUTION I begin by drawing a small sketch of the situation. Next to it I draw a circle with a cross to represent the center of mass of the book—no line to represent the floor, just the center-of-mass symbol representing the book as if it were by itself in space (**Figure 8.5**). Of the objects in the book's environment, the only one I need to consider is the floor because that is the only thing in contact with the book.

Figure 8.5



PROCEDURE: Drawing free-body diagrams

1. Draw a center-of-mass symbol (a circle with a cross) to indicate the object you wish to consider*—this object is your system. Pretend the object is by itself in empty space (hence the name *free body*). If you need to consider more than one object in order to solve a problem, draw a separate free-body diagram for each.
2. List all the items in the object's environment that are in contact with the object. These are the items that exert *contact forces* on the object. *Do not add these items to your drawing*! If you do, you run the risk of confusing the forces exerted *on* the object with those exerted *by* the object.
3. Identify all the forces exerted *on* the object by objects in its environment. (For now, omit from consideration any force not exerted along the object's line of motion.) In general, you should consider (*a*) the *gravitational field force* exerted by Earth on the object and (*b*) the *contact force* exerted by each item listed in step 2.
4. Draw an arrow to represent each force identified in step 3. Point the arrow in the direction in which the force is exerted and place the tail at the center of mass. If possible, draw the lengths of the arrows so that they reflect the relative magnitudes of the forces. Finally, label each arrow in the form

$$\vec{F}^{\text{type}}_{\text{by on}},$$

where "type" is a single letter identifying the origin of the force (c for contact force, *G* for gravitational force), "by" is a single letter identifying the object exerting the force, and "on" is a single letter identifying the object subjected to that force (this object is the one represented by the center of mass you drew in step 1).
5. Verify that all forces you have drawn are exerted **on** and not **by** the object under consideration. Because the first letter of the subscript on $\vec{F}$ represents the object exerting the force and the second letter represents the object on which the force is exerted, every force label in your free-body diagram should have the same second letter in its subscript.
6. Draw a vector representing the object's acceleration *next to* the center of mass that represents the object. Check that the vector sum of your force vectors points in the direction of the acceleration. If you cannot make your forces add up to give you an acceleration in the correct direction, verify that you drew the correct forces in step 4. If the object is not accelerating (that is, if it is in translational equilibrium), write $\vec{a} = \vec{0}$ and make sure your force arrows add up to zero. If you do not know the direction of acceleration, choose a tentative direction for the acceleration.
7. Draw a reference axis. If the object is accelerating, it is often convenient to point the positive x axis in the direction of the object's acceleration.

When your diagram is complete it should contain only the center-of-mass symbol, the forces exerted *on* the object (with their tails at the center of mass), an axis, and an indication of the acceleration of the object. Do not add anything else to your diagram.

*Representing, say, a car by a single point might seem like an oversimplification, but because we are interested only in the motion of the car as a whole, any details (like its shape or orientation) do not matter. Any unnecessary details you add to your drawing only distract from the issue at hand.

CONCEPTS

Next I name the forces exerted on the book. First there is a downward gravitational force exerted by Earth on the book. In addition, there is the one contact force exerted by the floor on the book. This force, which prevents the book from falling down through the floor, is directed upward.

Then I add arrows to the sketch to represent the forces I have identified. To represent the gravitational force, I draw an arrow pointing vertically downward with its tail at the center of mass. I label this force with a superscript *G* (for *gravity*) and the subscript Eb to indicate that it is a force exerted *by* Earth (E) *on* the book (b). For the contact force, I draw an upward-pointing arrow with its tail at the center of mass. I label this force with a superscript c (for *contact*) and the subscript fb to indicate that it is a force exerted *by* the floor (f) *on* the book (b). Finally, I verify that all forces in the diagram have subscripts ending in b because my diagram should contain only forces exerted *on* the book.

Because the book is at rest, its acceleration is zero, and therefore I write $\vec{a} = \vec{0}$ near the center of mass. From this information, I conclude that the vector sum of the forces exerted on the book (the sum of the two force vectors I have drawn) must be zero. I therefore adjust the lengths of the two arrows in the diagram so that they are of equal length, as shown in Figure 8.5. I complete the diagram by drawing an upward-pointing x axis. ✔

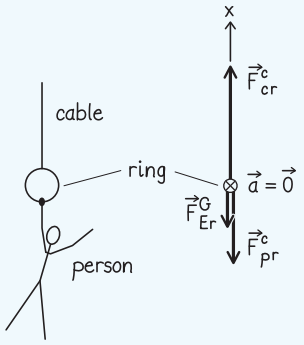
8.9 If Exercise 8.3 had asked about a book in free fall rather than one on the floor, what would the free-body diagram look like?

EXERCISE 8.4 Hanging around

Consider a person hanging motionless from a ring suspended from a cable, with the person's feet not touching the floor. Draw a free-body diagram for the ring.

SOLUTION I begin by drawing a sketch of the situation and then a center-of-mass symbol representing the ring (**Figure 8.6**). The things in the environment in contact with the ring are the cable and the person.

Figure 8.6



The next step is to identify all the forces exerted on the ring. As in the preceding example—and in *all* other examples concerning situations at or near Earth's surface—a vertically downward gravitational force is exerted by Earth (E) on the ring (r). I therefore draw an arrow from the center of mass down and label it $\vec{F}^{G}_{Er}$.

Next there are the contact forces: contact with the cable (c) and contact with the person (p). I represent the latter with a second arrow pointing downward. Because I don't know how this force compares in magnitude with the force of gravity, I choose its length arbitrarily. The other contact force is exerted by the cable on the ring. Because this force holds the ring up, I add an arrow pointing upward. Finally, I verify that the subscripts for each force end in r for ring because my diagram should contain only forces exerted *on* the ring.

Because the ring remains at rest, its acceleration is zero, and so I write $\vec{a} = \vec{0}$ near the center of mass. This tells me that the upward and downward forces exerted on the ring must add up to zero. I therefore adjust the length of the upward arrow so that it is equal to the sum of the lengths of the two downward arrows. I complete the diagram by adding an upward-pointing x axis. ✔

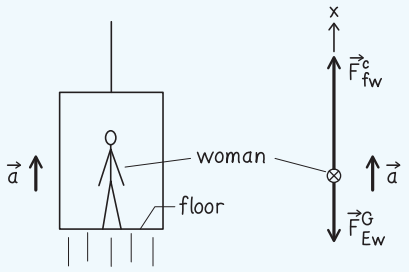
8.10 Draw a free-body diagram for the person in Exercise 8.4.

EXERCISE 8.5 Elevator woman

A woman stands in an elevator that is accelerating upward. Draw a free-body diagram for her.

SOLUTION After making a situation sketch, I draw a center-of-mass symbol to represent the woman (**Figure 8.7**). The only thing she is in contact with is the elevator floor—it is the floor that pushes her upward.

Figure 8.7



The forces exerted on the woman are a downward gravitational force exerted by Earth and an upward contact force exerted by the floor. I add arrows to the diagram to represent these forces, label them as shown in Figure 8.7, and verify that they all have subscripts ending in w for woman.*

What is the woman's acceleration? Even though she is at rest in the elevator, her acceleration is not zero because the elevator and she are both accelerating upward. I indicate this acceleration with an arrow pointing upward.

The upward acceleration means that the vector sum of the forces exerted on the woman is nonzero and directed upward. This can happen only if the upward force exerted by the elevator floor is larger than the gravitational force exerted by Earth. Therefore I adjust the lengths of the arrows as shown in Figure 8.7. Because the elevator has an upward acceleration, I add an upward-pointing x axis to my drawing. ✔

Warning: When drawing a free-body diagram, you may be tempted to add an arrow to show the vector sum of the forces (the "net force"), but this is not good practice. First, by drawing a vector showing the acceleration of the object, you have already indicated the direction of the vector sum of the forces. Second, this vector sum is the result of all the forces you have already drawn, not a separate force exerted by something on the object in your free-body diagram.

8.11 You throw a ball straight up. Draw a free-body diagram for the ball (*a*) while it is still touching your hand and is accelerating upward, (*b*) at its highest point, and (*c*) on the way back down.

*If all the subscripts in a free-body diagram do not end in the same letter, do not merely change the subscripts! Repeat the identification of the forces and ask yourself if each force you drew points in the correct direction. For example, does the floor push up on the woman? Or did you draw the force that her feet exert *on the elevator floor*? Correct your diagram if necessary.

CONCEPTS

8.6 Springs and tension

To better understand the origin of contact forces like the upward push of a floor or the pull of a rope, let us examine the behavior of a spring. When a spring is neither stretched nor compressed, we say it is at its *relaxed length*. If you position a spring vertically and put a brick on top of it, as shown in **Figure 8.8*a***, the brick *compresses* the spring and comes to rest in a position below the relaxed length of the spring. Figure 8.8 also shows the free-body diagram for the brick, which is subject to a downward force of gravity and an upward contact force—a *push*—exerted by the spring. Because the brick is at rest, the vector sum of the forces exerted on it must be zero. So the push by the spring and the gravitational force exerted by Earth are equal in magnitude and point in opposite directions. In general, whenever a spring holds an object in place, the spring must exert on the object an upward force, called a *support force*, that is equal in magnitude to the downward gravitational force exerted by Earth on the object.

Similarly, if we suspend the spring from a ceiling and fasten the brick at the bottom, the spring stretches, exerting on the brick an upward force—a *pull*—that counteracts the downward force of gravity (Figure 8.8*b*). The free-body diagram for this situation is identical to the one for the compressed spring; in both cases there is an upward force exerted by the spring on the brick. For the purpose of drawing a free-body diagram, it makes no difference at all whether it is a push or a pull that keeps the brick in place.

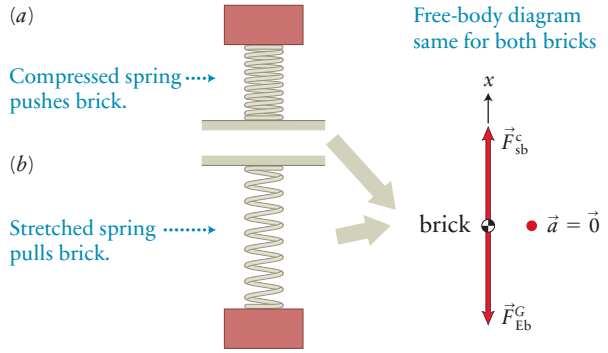
The force exerted by the spring always tends to return the spring to its relaxed length. When the spring is stretched, it exerts a pull (back toward the spring); when it is compressed, it exerts a push (away from the spring). Forces exerted by springs are therefore called *restoring forces*.

Now imagine adding additional identical bricks on top of the one shown in Figure 8.8*a*. Each additional brick increases the force exerted on the spring and further compresses it. If we repeat this experiment with the spring suspended from a ceiling, the spring stretches increasingly as the force exerted on the spring increases.

Over a certain range, called the *elastic range*, the compression or stretching of the spring is reversible; that is, when the force causing the compression or stretching is removed, the spring returns to its original shape.* If stretched beyond a certain point, called the *elastic limit*, the spring no longer reverts to its original length when the load (the brick) is removed. Additional load stretches the spring further until it eventually reaches the breaking point.

*As you might have noticed, the word *elastic* is associated with reversible phenomena: Elastic collisions are reversible, elastic potential energy is associated with reversible deformations, and so on.

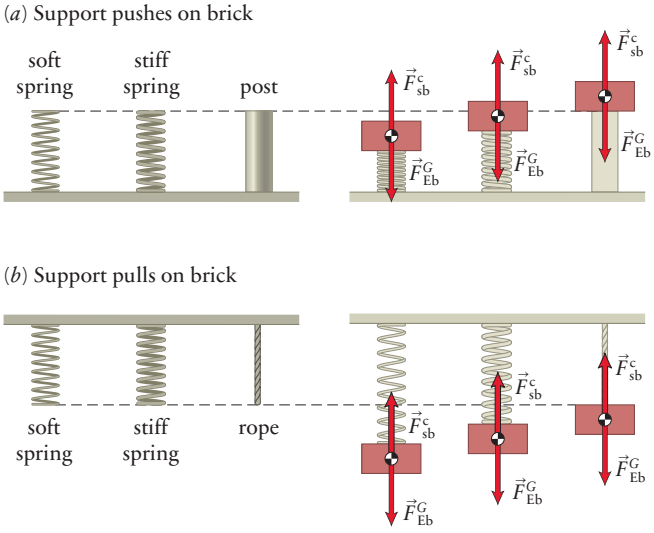
Figure 8.8 The free-body diagram for the brick is the same whether the spring (*a*) pushes or (*b*) pulls on it.



The amount of compression or stretching of a spring depends not only on the force exerted on the spring but also on the *stiffness* of the spring. A soft spring compresses or stretches more than a stiffer spring does under the same load (that is, when the same force is exerted on it). As **Figure 8.9** shows, however, the soft and stiff springs *exert exactly the same support force on the load*.

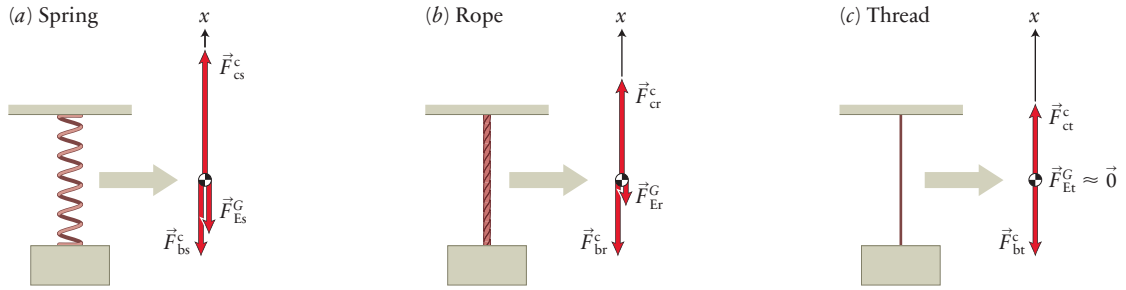
Now imagine putting a brick on a very stiff spring—one meant to support, say, a truck. You are unlikely to be able to see the spring compress (although you can measure the compression), but as with any other spring, compression is necessary in order to generate the upward force that holds the brick in place. With a very stiff spring, you need only a tiny compression to generate this support force. Even a solid object compresses a tiny bit when it supports another object, and as with a spring, the more load you pile on top of a solid object, the more it compresses in order to generate the force necessary to support its load.

Figure 8.9 A soft spring exerts the same force on a given load as a stiff spring or a post or a rope, even though it deforms much more. In fact, the post and rope behave like very stiff springs, compressing and stretching very slightly in response to the load.



CONCEPTS

Figure 8.10 Free-body diagrams for a spring, rope, and thread that suspend a brick from the ceiling.



Everything around you behaves this way: the chair you sit on, the bed you lie on, the floor you walk on, the road you drive on, and so on. Like a soft spring, an upholstered chair compresses noticeably when you sit on it. Sit on a table, though, and it "gives" a lot less. Even a concrete floor or a steel chair must compress ever so slightly to support you. The upholstered chair, the table, the concrete floor, and the steel chair all exert on you a support force that is equal in magnitude to the force of gravity exerted on you.

The same arguments can be applied to stretched objects (Figure 8.9*b*). In particular, the pull of a rope is like the pull of a stretched spring because the rope, like a very stiff spring, stretches ever so slightly while providing the support force that holds the object in place. As long as the deformation is reversible, the force exerted by a compressed or stretched material is called an **elastic force.**

Let us now turn our attention to the forces exerted *on* the object being compressed or stretched. **Figure 8.10*a*** shows a free-body diagram for a spring suspended from a ceiling and stretched by a brick. In addition to the two contact forces exerted by the ceiling and the brick on the spring, the spring is subject to a downward gravitational force. Because the spring is at rest, the upward force must be equal in magnitude to the sum of the downward forces exerted by gravity and the brick.

CONCEPTS

8.12 In Figure 8.10*a*, how does the magnitude of the downward force exerted *by the spring on the ceiling* compare with the magnitudes of the downward gravitational forces exerted by Earth on the spring and on the brick?

Check your answer to Checkpoint 8.12 before proceeding!

Figures 8.10*b* and 8.10*c* show free-body diagrams for a rope and a thread stretched by the same brick as in Figure 8.10*a*. The lower the inertia of the object being stretched, the smaller the force of gravity exerted on it. For objects of very low inertia, the magnitude of the gravitational force approaches zero, and the magnitudes of the two contact forces (exerted by the ceiling and the brick) become equal, as Figure 8.10*c* shows.

Because of the reciprocity of forces, the force exerted *by the thread on the ceiling* (not shown in the diagram) is equal in magnitude to the force exerted by the brick on the thread. Because these two forces point in the same direction, the thread simply transmits the pulling force of the brick to the ceiling.

Of course, we don't usually use thread to pull on objects, but in situations in which a rope—or even a lightweight chain—is used to pull, we can generally ignore the gravitational force exerted on the rope or chain. The only requirement is that the pulling force be much greater than the gravitational force exerted on the rope or chain. So, if we ignore the gravitational force, the same transmission of forces described for our thread is seen when we use a rope. If you fasten a rope to a boat and pull on one end of the rope, the rope stretches a little and the opposite end of the rope transmits your force to the boat. In general, if the force of gravity exerted on any object being stretched—rope, spring, thread—is much smaller than the forces that cause the stretching (and acceleration is zero), you can ignore the force of gravity. This is usually true when ropes, springs, or threads are used to pull or push on objects. Therefore:

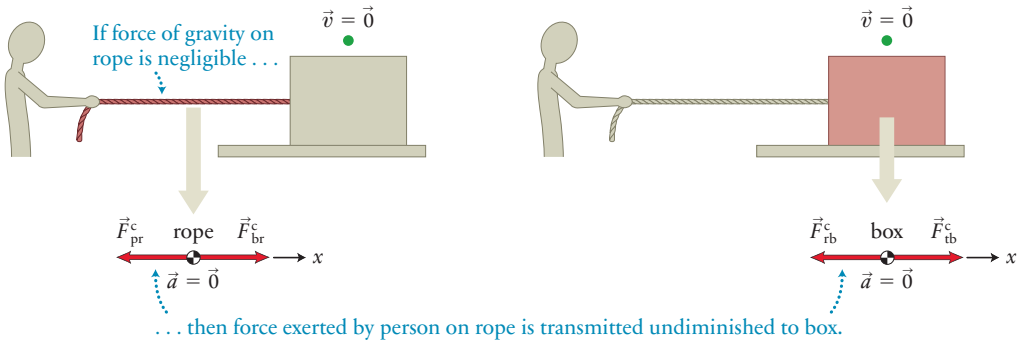
> **The force exerted on one end of a rope, spring, or thread is transmitted undiminished to the other end, provided the force of gravity on the rope, spring, or thread is much smaller than the forces that cause the stretching.**

A rope being used to pull an object is subject to two equal outward forces, one on each end (**Figure 8.11**). The stress caused by this pair of forces is called the **tension** in the rope, and the forces causing the tension are called *tensile forces*. Tension is a scalar and is represented by $\mathcal{T}$. In an object subject to two equal tensile forces of magnitude F, one on each end of the object, the tension is $\mathcal{T} = F$.

EXAMPLE 8.6 Tug of war

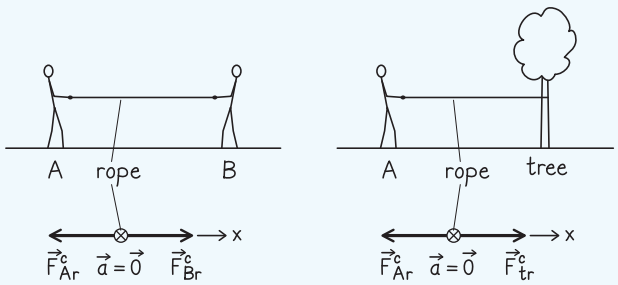
If two people, A and B, pull on opposite ends of a rope that is at rest, each exerting a horizontal tensile force of magnitude F, the tension in the rope is $\mathcal{T} = F$. Suppose instead that one end of the rope is tied to a tree and A pulls on the other end by himself with the same force magnitude F. Is the tension in the rope larger than, equal to, or smaller than the tension when A and B pull on opposite ends?

Figure 8.11 Transmission of a force by a rope to a box. Free-body diagrams for (*a*) the rope and (*b*) the box. (We assume that the force of gravity on the rope is much smaller than the forces exerted by the person and the box, so we can ignore it.)



1 GETTING STARTED To determine the tension in each case, I need to know the magnitude of the tensile force exerted on each end of the rope. I begin by making a sketch of the two situations (**Figure 8.12**).

Figure 8.12



2 DEVISE PLAN Because I must analyze various forces exerted on the rope, it is best to begin with a free-body diagram for the rope. In a tug of war only the horizontal forces matter, so I ignore the effect of the vertical gravitational force exerted by Earth and consider only the horizontal forces (Figure 8.12). Because the rope is always at rest, I know that the vector sum of the forces exerted on the rope must be zero. The magnitudes of the tensile forces exerted on either end are therefore equal to each other, and the tension in the rope is equal to the magnitude of the tensile force exerted on either end.

3 EXECUTE PLAN The rope tied to the tree is subject to two contact forces, one exerted by A and the other exerted by the tree. Because the vector sum of the forces exerted on the rope must be zero, I make the two force arrows of equal length, as shown in my free-body diagram.

Because A pulls on the rope with a force of magnitude F, the tree must pull with a force of magnitude F in the opposite direction, and so the tension in the rope is $\mathcal{T} = F$, which is the same as when A and B pull on opposite ends. ✔

4 EVALUATE RESULT The magnitude of the force exerted by A is F in both cases. Because the rope has zero acceleration in both cases, the magnitude of the force exerted by the tree must be the same as the magnitude of the force exerted by B when A and B pull. Therefore it makes sense that the tension in the rope is the same in the two cases.

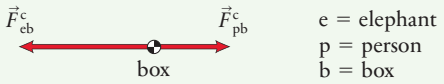
8.13 In Example 8.6, suppose both people pull on the same end of the rope, each exerting a force F, while the other end is still tied to the tree. Is the tension in the rope larger than, equal to, or smaller than the tension when the two people pull on opposite ends?

CONCEPTS

Self-quiz

1. (*a*) Two forces, one twice the magnitude of the other, are exerted on identical objects during the same time interval. Which force causes the greater change in momentum? (Assume that no other forces act on the objects.) (*b*) Two identical forces are exerted on identical objects, one during a time interval twice as long as the other. Which force causes the greater change in momentum? (Assume that no other forces act on the objects.)
2. The free-body diagram of a box subjected to forces is shown in Figure 8.13. Because we are concerned about only the horizontal motion, the forces in the vertical direction have been omitted. (*a*) What is the direction of the change in the momentum of the box? (*b*) If the box is at rest before the forces shown in the diagram are exerted on it, what is the direction of its momentum once the forces are exerted?

Figure 8.13



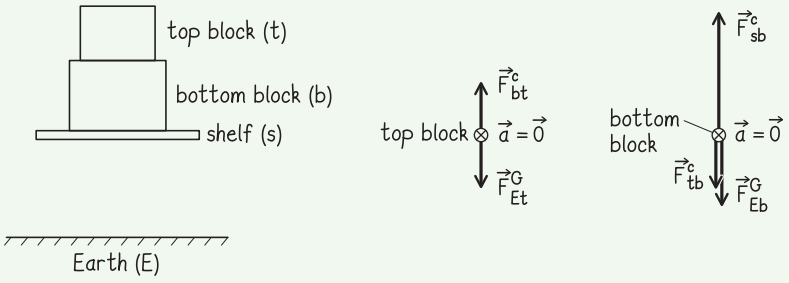
3. For the free-body diagram in question 2, list the objects in the environment of the box.
4. A block of wood rests on a shelf, and a second block of wood rests on top of the first. (*a*) List the objects in the environment of the top block. (*b*) List the objects in the environment of the bottom block. (*c*) Draw a free-body diagram for each block.
5. Complete the following sentences, applying the reciprocity of forces. (Example: If I push down on the seat of my chair, *the seat of my chair pushes up on me*.)

   If I push to the right on the wall, . . .
   If I pull to the right on a spring, . . .
   If the floor pushes up on my feet, . . .
   If Earth pulls down on me, . . .

Answers

1. (*a*) The larger force causes the greater change in momentum because force and change in momentum are proportional to each other over a given time interval. (*b*) The force exerted over the longer time interval causes the greater change in momentum because time interval and change in momentum are proportional to each other for a constant force.
2. (*a*) The vector sum of the forces exerted on the box points to the left because the force exerted on the box by the elephant is larger than the force exerted on the box by the person. The change in momentum is to the left, in the same direction as the vector sum of the forces. (*b*) The change in momentum is equal to the final momentum minus the initial momentum of zero. Therefore, the final momentum of the box must be to the left.
3. The objects in the environment of the box are an elephant and a person. If you also consider the vertical direction, you need to include the ground and Earth as being in the environment of the box.
4. (*a*) The objects in the environment of the top block are the bottom block and Earth. The shelf is not in contact with the top block and so is not part of the top block's environment. (*b*) The objects in the environment of the bottom block are the shelf, the top block, and Earth. (*c*) See Figure 8.14. Earth pulls downward on each block. The surface below each block pushes upward. If the bottom block pushes upward on the top block, the top block must push down on the bottom block.

Figure 8.14



5. For each, swap the nouns and reverse the given direction.

   If I push to the right on the wall, *the wall pushes to the left on me.*
   If I pull to the right on a spring, *the spring pulls to the left on me.*
   If the floor pushes up on my feet, *my feet push down on the floor.*
   If Earth pulls down on me, *I pull up on Earth.**

*Many people find this last statement difficult to accept and are inclined to answer: If Earth pulls down on me, the ground pushes up on me. That answer, however, includes three objects (ground, Earth, me), not just the two required by the reciprocity of forces.

8.7 Equation of motion

In the first part of this chapter, we determined that the vector sum of the forces exerted on an object is equal to the time rate of change in the object's momentum. If an object's momentum changes by an amount $\Delta\vec{p}_{\text{object}}$ in a time interval Δt, the *average* time rate of change in the object's momentum is

$$\frac{\Delta\vec{p}_{\text{object}}}{\Delta t}. \tag{8.1}$$

The *instantaneous* time rate of change in momentum is obtained by letting the time interval approach zero:

$$\lim_{\Delta t\to 0}\frac{\Delta\vec{p}_{\text{object}}}{\Delta t} = \frac{d\vec{p}_{\text{object}}}{dt} \equiv \Sigma\vec{F}_{\text{object}}, \tag{8.2}$$

where $\Sigma\vec{F}_{\text{object}}$ represents the vector sum of all the forces exerted on the object:*

$$\Sigma\vec{F}_{\text{object}} = \vec{F}_{1\text{ object}} + \vec{F}_{2\text{ object}} + \vec{F}_{3\text{ object}} + \cdots, \tag{8.3}$$

where $\vec{F}_{1\text{ object}}$ is the force exerted by object 1 on the object under consideration, and so on. Because we are concerned with only one object, we can omit the subscript on the left side of Eq. 8.3, and so Eq. 8.2 becomes

$$\Sigma\vec{F} \equiv \frac{d\vec{p}}{dt}. \tag{8.4}$$

The vector sum of the forces points in the same direction as the change in momentum of the object.

Substituting $\vec{p} \equiv m\vec{v}$ in the term on the right in Eq. 8.4 gives

$$\frac{d\vec{p}}{dt} \equiv \frac{d(m\vec{v})}{dt} = m\frac{d\vec{v}}{dt}, \tag{8.5}$$

where I have used the fact that the inertia m of the object is constant. Because $d\vec{v}/dt$ is the acceleration $\vec{a}$ of the object, substituting Eq. 8.5 into Eq. 8.4 yields

$$\Sigma\vec{F} = m\vec{a} \tag{8.6}$$

or

$$\vec{a} = \frac{\Sigma\vec{F}}{m}. \tag{8.7}$$

In terms of components, Eqs. 8.6 and 8.7 become

$$\Sigma F_x = ma_x \quad \text{and} \quad a_x = \frac{\Sigma F_x}{m}. \tag{8.8}$$

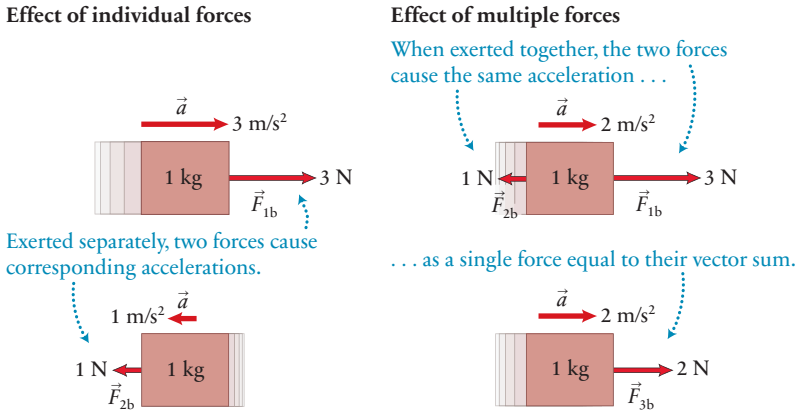
Equation 8.7 is called the **equation of motion** for the object. For a given vector sum of forces, this equation allows us to obtain the acceleration and therefore to determine the motion of the object. Conversely, if we know the acceleration of an object, Eq. 8.6 allows us to determine the vector sum of the forces exerted on that object.

Because the SI unit of momentum is kg · m/s and the SI unit of time is seconds, the SI unit of force is kg · m/s^2. This derived SI unit is given the special name **newton** (N) in honor of the English scientist Isaac Newton:

$$1\text{ N} \equiv 1\text{ kg}\cdot\text{m/s}^2. \tag{8.9}$$

*We drop the "by" subscript because the sum represents the forces exerted by *all* objects on the object under consideration.

Figure 8.15 Superposition of forces.



This definition means that a force of 1 N exerted on an object that has an inertia of 1 kg causes that object to accelerate at 1 m/s^2.

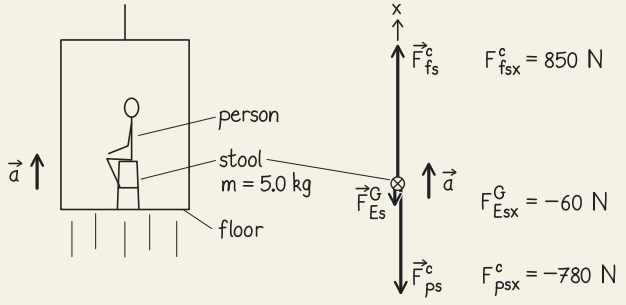
Note that while several forces can be exerted simultaneously on a single object, that object can have only one acceleration. For example, if a 1-kg object is simultaneously subject to two forces—say, a 3-N force directed to the right and a 1-N force directed to the left—the object accelerates at 2 m/s^2 to the right, just as if a single force of 2 N directed to the right had been exerted (see **Figure 8.15**). The reason forces can be added vectorially is that the effect of a force is a change in momentum, and momentum is a vector quantity. This property of adding forces vectorially is called the *superposition of forces*.

EXAMPLE 8.7 Elevator stool

A person is sitting on a stool in an elevator. The forces exerted on the stool are a downward force of magnitude 60 N exerted by Earth, a downward force of magnitude 780 N exerted by the person, and an upward force of magnitude 850 N exerted by the elevator floor. If the inertia of the stool is 5.0 kg, what is the acceleration of the elevator?

1 GETTING STARTED I know all the forces that are exerted on the stool, and I must determine its acceleration. I therefore begin by drawing a free-body diagram for the stool that includes the three forces mentioned in the problem statement (**Figure 8.16**). I arbitrarily choose an x axis pointing upward.

Figure 8.16



2 DEVISE PLAN Equation 8.8 gives a relationship between the vector sum of the forces exerted on an object and the acceleration of the object. By substituting the forces given and the inertia of the stool, I can solve for a_x.

3 EXECUTE PLAN From Eq. 8.8, I have

$$a_x = \frac{\Sigma F_x}{m} = \frac{1}{m}(F^c_{fsx} + F^c_{psx} + F^G_{Esx})$$

$$= \frac{1}{5.0\ \text{kg}}[(+850\ \text{N}) + (-780\ \text{N}) + (-60\ \text{N})]$$

$$= +2.0\ \text{m/s}^2. \checkmark$$

4 EVALUATE RESULT The acceleration I obtained is positive, which means that the elevator has an upward acceleration, which is what I expect because the force exerted by the floor on the stool is upward. The magnitude of the acceleration is about one-fifth the acceleration due to gravity, which strikes me as a reasonable value.

8.14 (*a*) You exert a constant force of 200 N on a friend on roller skates. If she starts from rest, estimate how far she moves in 2.0 s. (*b*) When a person jumps off a wall, what is the magnitude of his acceleration? (*c*) Estimate the magnitude of the force exerted by Earth on the person during the jump in part *b*.

QUANTITATIVE TOOLS

As we saw in Section 8.2, forces always come in interaction pairs. To see this quantitatively, consider an isolated system of two colliding carts. As we have seen before, the momentum of the two-cart system is constant, and so the changes in momentum caused by the collision must obey the relationship

$$\Delta\vec{p}_1 + \Delta\vec{p}_2 = \vec{0}, \tag{8.10}$$

where $\Delta\vec{p}_1$ is the change in momentum of cart 1 and $\Delta\vec{p}_2$ is that of cart 2. During an infinitesimal time interval dt, this relationship yields

$$d\vec{p}_1 + d\vec{p}_2 = \vec{0}. \tag{8.11}$$

Because dt is the same for both carts, the time rates of change in momentum must also add to zero:

$$\frac{d\vec{p}_1}{dt} + \frac{d\vec{p}_2}{dt} = \vec{0}. \tag{8.12}$$

According to Eq. 8.4, the time rate of change of momentum of an object is equal to the vector sum of the forces exerted on that object, and so

$$\Sigma\vec{F}_1 + \Sigma\vec{F}_2 = \vec{0}. \tag{8.13}$$

Because the carts are constrained to move in the horizontal direction, the forces in the vertical direction cancel, and so the only forces exerted on the carts we need to consider are those they exert on each other. The vector sum of the forces exerted on cart 1 is thus just the force exerted by cart 2 on cart 1, and the vector sum of the forces exerted on cart 2 is just the force exerted by cart 1 on cart 2:

$$\Sigma\vec{F}_1 = \vec{F}_{21} \quad \text{and} \quad \Sigma\vec{F}_2 = \vec{F}_{12}, \tag{8.14}$$

and thus $\vec{F}_{12} + \vec{F}_{21} = \vec{0}$, or

$$\vec{F}_{12} = -\vec{F}_{21}. \tag{8.15}$$

Newton's laws of motion

In the modern view, the conservation laws are the heart and soul of physics—they best represent the generally held belief that the universe has an underlying simplicity. In the 17th century, however, 200 years before the concept of energy and the conservation laws were formulated, the English scientist Isaac Newton published a work describing his ideas concerning the fundamental principles that govern forces and their effects on the motion of objects. Newton's three laws of motion, as they are known, have been the backbone of physics for more than two centuries.

We have already discussed each of these three laws—albeit under different, and in the current context more appropriate, names.

Newton's first law of motion, first formulated by Galileo Galilei, is what we called the *law of inertia* in Chapter 6:

> **In an inertial reference frame, any isolated object that is at rest remains at rest, and any isolated object that is in motion keeps moving at a constant velocity.**

Newton's second law of motion corresponds to the *definition of force* given in Eq. 8.4:

> **The vector sum of the forces exerted on an object is equal to the time rate of change in the momentum of that object.**

Equation 8.6, $\Sigma\vec{F} = m\vec{a}$, is frequently referred to as Newton's second law even though it holds only when m is constant.

Newton's third law of motion expresses the law of conservation of momentum in terms of forces (Eq. 8.15):

> **Whenever two objects interact, they exert on each other forces that are equal in magnitude but opposite in direction.**

Newton is widely regarded as one of the greatest physicists of all time, not only for his formulation of the laws of motion but also for his law of gravity, for his development of the calculus, and for laying the foundation of modern optics.

In words, whenever two objects interact, they exert forces on each other that are equal in magnitude and opposite in direction. The forces $\vec{F}_{12}$ and $\vec{F}_{21}$ form an **interaction pair.** Because interactions always involve two objects, every force that is exerted *on* an object has associated with it a force of equal magnitude in the opposite direction exerted *by* the object. Note, in Eq. 8.15, that the subscripts on the forces that form an interaction pair are interchanged. For that reason, both forces in an interaction pair can never be part of the same free-body diagram. Equation 8.15 is called **Newton's third law of motion** (see the box "Newton's laws of motion" on page 261); it is a direct consequence of the conservation of momentum expressed in Eq. 8.10.

8.15 If forces always come in interaction pairs and the forces in such a pair are equal in magnitude and opposite in direction (Eq. 8.15), how can the vector sum of the forces exerted on an object ever be nonzero?

8.8 Force of gravity

We have seen before that, when we ignore air resistance, all objects dropped from a certain height near Earth's surface have a downward acceleration given by $a_x = -g$ when we define up as being the positive x direction (see Chapter 3). For an object in free fall near Earth's surface, the only force exerted on the object is the gravitational force $\vec{F}^G_{\text{Eo}}$ exerted by Earth on the object. The left side of Eq. 8.8 is thus

$$\Sigma F_x = F^G_{\text{Eo}x}. \tag{8.16}$$

The right side is $ma_x = -mg$, where m is the inertia of the object, and so the x component of the gravitational force exerted by Earth on the object is

$$F^G_{\text{Eo}x} = -mg \text{ (near Earth's surface, } +x \text{ axis vertically upward).} \tag{8.17}$$

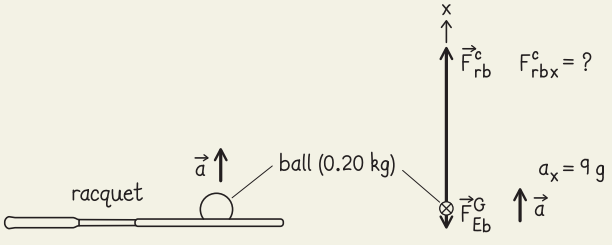
The minus sign reflects the fact that the acceleration due to gravity (and therefore also the force of gravity) is directed downward and the positive x direction points upward. Because the acceleration due to gravity is the same for all objects, Eq. 8.17 holds for all objects in free fall near Earth's surface. Note that when additional forces are exerted on the object, the x component of the gravitational force exerted by Earth on the object is still given by Eq. 8.17, even though the x component of the object's acceleration is no longer equal to $-g$.

EXAMPLE 8.8 Tennis ball launch

A tennis ball of inertia 0.20 kg is launched straight up in the air by hitting it with a racquet. If the magnitude of the acceleration of the ball while it is in contact with the racquet is $9g$, what are the magnitude and direction of the force exerted by the racquet on the ball?

❶ GETTING STARTED I am given the inertia and acceleration magnitude for a tennis ball hit by a racquet, and I must calculate the force—both magnitude and direction—responsible for this acceleration. Because this problem deals with forces exerted on the ball, I begin by drawing a situation sketch and a free-body diagram for the ball (Figure 8.17). While it is being hit, the ball is in contact with the racquet and so I include an upward contact force exerted by the racquet as well as a downward force of gravity exerted by Earth. Because the ball accelerates upward, I make the upward force vector longer than the downward force vector and point the positive x axis upward.

Figure 8.17



❷ DEVISE PLAN Equation 8.8 relates the ball's acceleration to the vector sum of the two forces exerted on it. I can use Eq. 8.17 to calculate one of the two forces and then solve for the other one in Eq. 8.8.

QUANTITATIVE TOOLS

3 EXECUTE PLAN Equation 8.8 yields

$$\Sigma F_x = F^G_{\mathrm{Eb}x} + F^c_{\mathrm{rb}x} = ma_x.$$

Because I've defined upward as the positive x direction, $a_x = +9g$. Substituting this value and Eq. 8.17 for the gravitational force exerted by Earth on the object, I get

$$(-mg) + F^c_{\mathrm{rb}x} = +9mg$$

or
$$F^c_{\mathrm{rb}x} = +10mg = +10\,(0.20\ \mathrm{kg})\,(9.8\ \mathrm{m/s^2})$$
$$= +20\ \mathrm{N}. \checkmark$$

4 EVALUATE RESULT The plus sign indicates that this force is directed upward, as expected. The magnitude of the force is ten times larger than the gravitational force exerted by Earth, which also makes sense if the ball is to be accelerated upward during the very short time interval that the racquet is in contact with the ball.



8.16 The magnitude of the gravitational force exerted by Earth on an object of inertia m_1 is $m_1 g$. (*a*) What is the magnitude of the force exerted by the object on Earth (inertia m_E)? (*b*) What is the acceleration of Earth due to its gravitational interaction with the object?

Equation 8.17 shows that the magnitude of the force of gravity exerted on an object is proportional to the object's inertia, which is an intrinsic property of the object that has nothing to do with the presence of Earth or with gravity. This surprising conclusion is a direct consequence of the fact that, when air resistance is negligible, all objects experience the same acceleration due to gravity. It is more difficult to accelerate a large boulder than a feather, but the gravitational forces exerted by Earth on the boulder and on the feather are such that both have the same acceleration when they fall (ignoring air resistance). Equation 8.17 doesn't tell us much about the mechanism of gravity; it merely tells us what magnitude and direction the force of gravity must have to account properly for the free fall of objects near the surface of Earth. In Chapter 13, we'll return to the subject of gravity and its relationship to inertia. In the meantime, we can use Eq. 8.17 to calculate the magnitude of the force of gravity on any object near the surface of Earth.



8.17 Suppose you are in an elevator that is accelerating upward at 1 m/s^2. (*a*) Draw a free-body diagram for your body. (*b*) Determine the magnitude of the force exerted by the elevator floor on you.

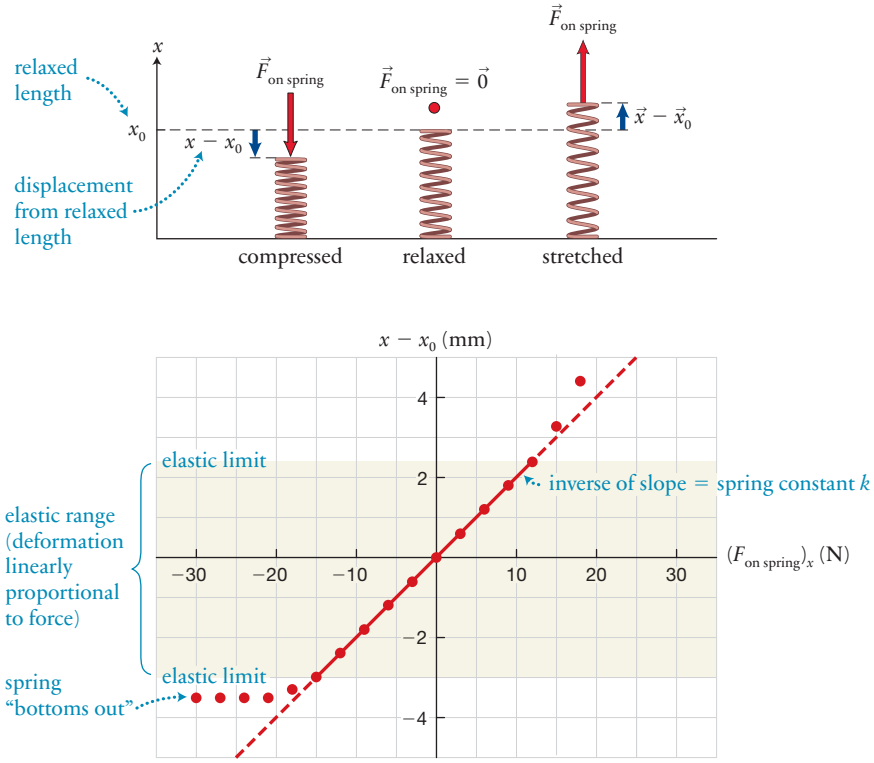
8.9 Hooke's law

Now that we have established the relationship between the inertia of an object and the gravitational force exerted on it (Eq. 8.17), we can analyze the behavior of springs quantitatively. We begin by attaching bricks to the free end of a spring that is suspended from a ceiling. Each additional brick increases the force exerted on the spring and the amount by which the spring stretches. We repeat this experiment, but now we support the spring on a solid surface and stack bricks on top of it. Each additional brick increases the force exerted on the spring and further compresses it, up to the point where the spring "bottoms out"—which means its coils get in the way of further compression.

The graph in **Figure 8.18** on the next page shows the x component of the displacement of the free end of the spring from its relaxed position at x_0 versus the x component of the force exerted *on* the spring *by* the bricks. The shaded region indicates the elastic range. The first thing to note is that, within the elastic range, the displacement and the force exerted by the load on the spring are linearly proportional to each other. The second thing to note is that the force exerted on the spring is always in the same direction as the displacement. We can express this quantitatively as

$$(F_{\text{by load on spring}})_x = k(x - x_0) \quad \text{(small displacement)}, \qquad (8.18)$$

Figure 8.18 The x component of the displacement of the free end of the spring from its relaxed position is proportional to the x component of the force causing the displacement.



where k is called the **spring constant.** The spring constant represents the inverse of the slope of the straight line through the data points in Figure 8.18; it tells us how large a force is required to give the free end of the spring a displacement $x - x_0$. The spring constant is therefore a quantitative measure of the stiffness of the spring discussed in Section 8.6. It is always positive, but different springs have different values of k, depending on how they respond to compression and stretching. The SI unit of the spring constant is N/m.

Often we are interested not in the force exerted *on* the spring, but in the force exerted *by* the spring. We know from Eq. 8.15 that

$$\vec{F}_{\text{by spring on load}} = -\vec{F}_{\text{by load on spring}} \tag{8.19}$$

and so the force exerted by the spring is given by

$$(F_{\text{by spring on load}})_x = -k(x - x_0) \quad \text{(small displacement)}. \tag{8.20}$$

Note that the direction of the force of the spring on the load is opposite to the direction of the displacement. A stretched spring pulls back. A compressed spring pushes out. Equation 8.18 (or its equivalent, Eq. 8.20) is called **Hooke's law** after the British physicist Robert Hooke (1635–1703), who first discovered the linear relationship between a spring's stretching or compression and the force exerted on it.

Unlike the conservation principles, which are of universal validity, Eq. 8.20 is valid for only small displacement. As Figure 8.18 shows, Hooke's "law" fails if we compress or stretch the spring beyond certain limits. Still, Hooke's law is of great use because, for small deformations, many materials—from steel to foam—show a linear relationship between force and displacement when compressed or stretched. For this reason, the force that tends to return a stretched or compressed spring or material to its relaxed position is referred to as a *linear restoring force*.

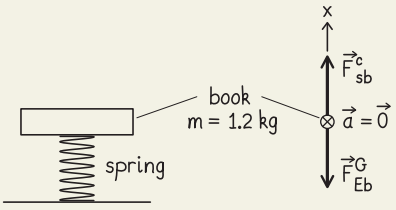
QUANTITATIVE TOOLS

EXAMPLE 8.9 Spring compression

A book of inertia 1.2 kg is placed on top of the spring in Figure 8.18. What is the displacement of the top end of the spring from the relaxed position when the book is at rest on top of the spring?

❶ **GETTING STARTED** My task is to determine how much the top end of a spring is displaced when a 1.2-kg book is placed on the spring. Hooke's law tells me that this displacement is proportional to the force exerted by the spring on the book, so after making a sketch of the situation, I draw a free-body diagram for the book. The book is subject to a downward gravitational force exerted by Earth and an upward contact force exerted by the spring (**Figure 8.19**). Because the book is in translational equilibrium, the book's acceleration is zero and so the two forces are equal in magnitude. I arbitrarily point the x axis upward.

Figure 8.19



❷ **DEVISE PLAN** I can use Eq. 8.20 to determine the displacement, but in order to do so I need to know the spring constant k and the x component F^{c}_{sbx} of the force exerted by the spring on the book. The spring constant is given by the slope of the curve in Figure 8.18. From my free-body diagram I know that $\vec{F}^{c}_{sb}$ is equal to the negative of the force of gravity $\vec{F}^{G}_{Eb}$ exerted by Earth on the book.

❸ **EXECUTE PLAN** From the graph in Figure 8.18, I arbitrarily choose 25 N as my force value and see that it causes a displacement of 5.0 mm. The spring constant is thus, from Eq. 8.18:

$$k = \frac{(F_{\text{by load on spring}})_x}{x - x_0} = \frac{+25\ \text{N}}{+5.0 \times 10^{-3}\ \text{m}} = 5.0 \times 10^{3}\ \text{N/m}.$$

Because the book is at rest, I know that $\vec{F}^{c}_{sb} + \vec{F}^{G}_{Eb} = \vec{0}$, and so $\vec{F}^{G}_{Eb} = -\vec{F}^{c}_{sb}$. The force of gravity is given by Eq. 8.17, and so I get for the force exerted by the spring on the book

$$\begin{aligned} F^{c}_{sbx} &= -F^{G}_{Ebx} \\ &= -(-mg) \\ &= (1.2\ \text{kg})(9.8\ \text{m/s}^2) \\ &= +12\ \text{N}. \end{aligned}$$

Substituting this value and the spring constant into Eq. 8.20, I obtain for the displacement:

$$\begin{aligned} x - x_0 &= \frac{-F^{c}_{sbx}}{k} \\ &= \frac{-12\ \text{N}}{5.0 \times 10^{3}\ \text{N/m}} \\ &= -2.4 \times 10^{-3}\ \text{m}. \checkmark \end{aligned}$$

❹ **EVALUATE RESULT** It makes sense that the force exerted by the book on the spring is equal to the force exerted by Earth on the book ($\vec{F}^{G}_{Eb} = \vec{F}^{c}_{bs}$) because the spring has to keep the book from being pulled downward into the ground. I can thus also obtain my answer by reading off the displacement from Figure 8.18 for a force of magnitude 12 N (the magnitude of the force of gravity exerted on the book). This procedure yields a value that is very close to the result I obtained. The negative sign in my answer indicates that the top end of the spring is displaced downward, as I expect.



8.18 (*a*) Is a spring that has a large spring constant k stiffer or softer than a spring that has a small spring constant? (*b*) Which has a larger spring constant: steel or foam rubber?

8.10 Impulse

Back in Chapter 4 we defined the change in momentum of a system or object as the impulse delivered to it. In this section we determine a relationship between force and impulse, which permits us to determine the effect forces have on momentum.

We begin by considering a simple case: a single object subject to constant forces. To calculate what happens to the momentum of the object, we consider its motion during a time interval $\Delta t = t_f - t_i$. The constant

forces exerted on the object give it a constant acceleration $\vec{a}$ given by Eq. 8.7. Because the acceleration is constant, we can write

$$\vec{a} = \frac{\Delta\vec{v}}{\Delta t}. \tag{8.21}$$

Multiplying both sides by the inertia m of the object, we get

$$m\vec{a} = m\frac{\Delta\vec{v}}{\Delta t} = \frac{\Delta\vec{p}}{\Delta t} \tag{8.22}$$

or

$$\Delta\vec{p} = m\vec{a}\,\Delta t. \tag{8.23}$$

From Eq. 8.6 we know that $\Sigma\vec{F} = m\vec{a}$, and so Eq. 8.23 becomes

$$\Delta\vec{p} = \left(\Sigma\vec{F}\right)\Delta t \quad \text{(constant force)}. \tag{8.24}$$

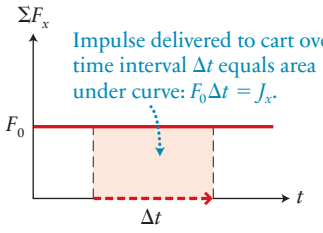
If we compare this to the momentum law ($\Delta\vec{p} = \vec{J}$, Eq. 4.18), we see that the impulse delivered by the forces to the object is

$$\vec{J} = \left(\Sigma\vec{F}\right)\Delta t \quad \text{(constant force)}. \tag{8.25}$$

This equation, called the **impulse equation,** states that the impulse delivered to an object—the change in an object's momentum—during a time interval Δt is equal to the product of the vector sum of the forces exerted on the object and the duration of the time interval. Equation 8.24 shows that the longer a force is exerted, the smaller it has to be to accomplish a given change in the momentum of the object. This is why, in the bicycling example discussed in Section 8.1, the force of impact is less when you hit the mattress instead of the concrete wall. Likewise, it hurts less to catch a ball (that is to say, bring the ball to rest) with a padded glove than without one. And, by holding out your hand in front of you and then drawing it back just as the ball strikes your glove, you can further increase Δt and thus reduce the force required to slow the ball to a stop.

Figure 8.20 The impulse delivered by a force is equal to the area under the $F_x(t)$ curve, as shown here for (*a*) a constant force and (*b*) a force that varies with time.

(*a*) Impulse delivered by constant force exerted on object



(*b*) Hand shoves cart, exerting force that varies with time

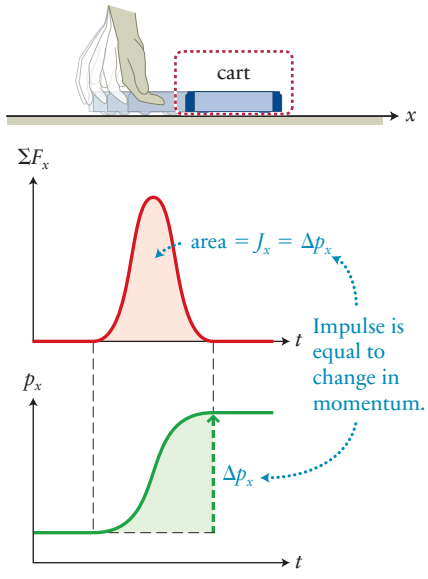


Figure 8.20*a* shows a force-versus-time graph for an object subject to constant forces whose vector sum is equal to F_0. The impulse delivered during a time interval Δt is equal to the shaded rectangular area under the $F_x(t)$ curve—the width Δt times the height F_0.

So far in this discussion, the impulse-delivering forces we've looked at have been *constant* forces. However, the relationship between impulse and area under the $F_x(t)$ curve holds not only when $\Sigma\vec{F}$ is constant. Just as we found, in Section 3.8, that the area under the $a_x(t)$ curve of an object yields the change in velocity (Eq. 3.27, $\Delta v_x = \int_{t_i}^{t_f} a_x(t)dt$), we calculate for the impulse delivered by an arbitrary force $\Sigma\vec{F}$ that varies with time

$$\Delta\vec{p} = \vec{J} = \int_{t_i}^{t_f} \Sigma\vec{F}(t)dt \quad \text{(time-varying force)}. \tag{8.26}$$

You can obtain this equation by integrating Eq. 8.4. You can also convince yourself of this result either by applying to force the procedure we applied to acceleration in Section 3.8 or by multiplying the left and right sides of Eq. 3.27 by the inertia m of the object.

Figure 8.20*b* shows the momentum of a cart getting a shove. The impulse delivered by the hand to the cart is equal to the change in momentum $\Delta\vec{p}$ shown in the figure. The force exerted by the hand on the cart is also shown in Figure 8.20*b*. This curve is obtained by applying what we know from Eq. 8.4 and taking the derivative of the $p_x(t)$ curve in Figure 8.20*b*. As the hand begins pushing, the force increases and peaks. Then, as the

QUANTITATIVE TOOLS

hand stops pushing, the force decreases to zero. The impact of a bat on a ball varies in a similar way with time, albeit much more rapidly (Figure 8.21). In general, the impulse delivered by a time-varying force during a given time interval is equal to the area under the $F_x(t)$ curve during that time interval (Eq. 8.26 and Figure 8.20*b*).

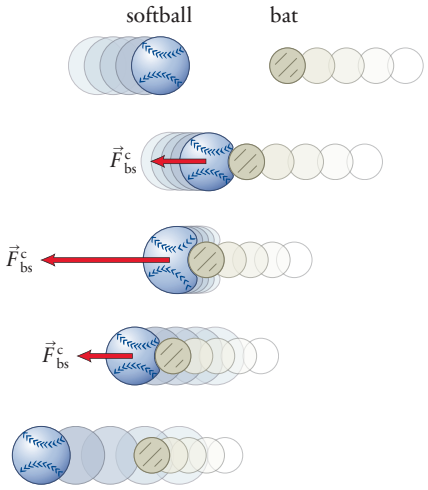
If the vector sum of the forces exerted on the object is zero, $\Sigma\vec{F} = \vec{0}$, the object is in translational equilibrium. Equation 8.26 then tells us that the momentum of the object does not change

$$\Delta\vec{p} = \vec{0} \quad \text{(translational equilibrium)}. \tag{8.27}$$

In words, the momentum of an object in translational equilibrium does not change.

8.19 (*a*) A feather and a brick are falling freely in an evacuated tube. Is the magnitude of the gravitational force exerted by Earth on the feather larger than, smaller than, or equal to that exerted by Earth on the brick? (*b*) Suppose equal forces of 10 N are exerted on both objects for 2 s. Which object gains more momentum? (*Evacuated* means all the air has been sucked out of the tube, so that there is no air resistance to hinder the motion.)

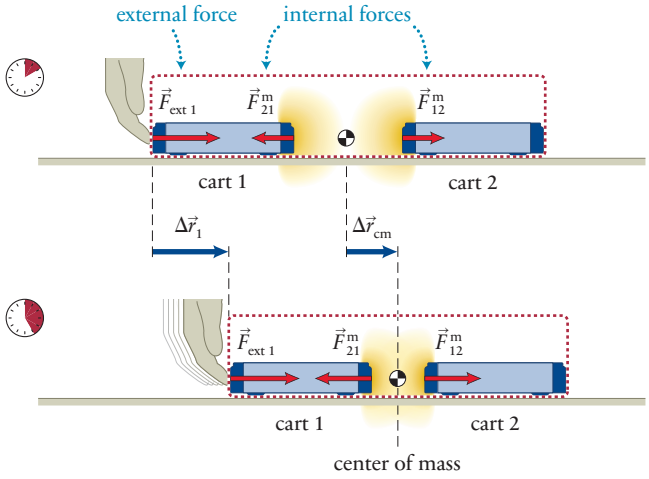
Figure 8.21 As a bat strikes a ball, the magnitude of the force exerted by the bat on the ball increases as the ball deforms, peaks at the instant the ball and bat have zero relative velocity, and finally decreases back to zero as the ball regains its shape and leaves the bat.



8.11 Systems of two interacting objects

So far we have considered only the very simple case of a single object subject to a force. In this section we look at a system of two interacting objects. For such systems, it is useful to make a distinction between forces exerted from inside the system and forces exerted on the system from outside. The former are called *internal forces*, the latter *external forces*. In Figure 8.22, two carts equipped with repelling magnets move on a track while cart 1 gets a push. Because of the magnetic repulsion, the carts never touch each other; instead, cart 2 accelerates to the right when cart 1 approaches. If we choose the two carts as our system, the forces the carts exert on each other are internal, and the push exerted on cart 1 is external.

Figure 8.22 Two carts with repelling magnets moving under the influence of an external force exerted on cart 1.



The momentum of the two-cart system is

$$\vec{p} = \vec{p}_1 + \vec{p}_2. \tag{8.28}$$

Differentiating this equation with respect to time and substituting the vector sum of the forces for each dp/dt term (Eq. 8.4) give

$$\frac{d\vec{p}}{dt} = \frac{d\vec{p}_1}{dt} + \frac{d\vec{p}_2}{dt} = \Sigma\vec{F}_1 + \Sigma\vec{F}_2. \tag{8.29}$$

The vector sum of the forces exerted on cart 1 consists of two parts: the external force exerted on cart 1 and an internal force exerted by cart 2 on cart 1 (due to the magnetic repulsion):

$$\Sigma\vec{F}_1 = \vec{F}_{\text{ext }1} + \vec{F}^{\text{m}}_{21}. \tag{8.30}$$

Because no external force is exerted on cart 2, the vector sum of the forces exerted on it is

$$\Sigma\vec{F}_2 = \vec{F}^{\text{m}}_{12}. \tag{8.31}$$

Substituting Eqs. 8.30 and 8.31 into the right side of Eq. 8.29 gives

$$\frac{d\vec{p}}{dt} = \vec{F}_{\text{ext }1} + \vec{F}^{\text{m}}_{21} + \vec{F}^{\text{m}}_{12}. \tag{8.32}$$

Because $\vec{F}^{\text{m}}_{12}$ and $\vec{F}^{\text{m}}_{21}$ form an interaction pair, we have $\vec{F}^{\text{m}}_{12} = -\vec{F}^{\text{m}}_{21}$, and so

$$\frac{d\vec{p}}{dt} = \vec{F}_{\text{ext }1}. \tag{8.33}$$

Equation 8.33 shows that only the external force exerted on cart 1 changes the momentum of the system. The internal forces do not have any effect on the momentum of the system.

Using the definition of center-of-mass velocity from Section 6.6, we have $\vec{p} = m\vec{v}_{\text{cm}}$, with $m = m_1 + m_2$. Because the inertia m of the system remains constant in the interaction, we have

$$\frac{d\vec{p}}{dt} = \frac{d(m\vec{v}_{\text{cm}})}{dt} = m\frac{d\vec{v}_{\text{cm}}}{dt} \equiv m\vec{a}_{\text{cm}} \tag{8.34}$$

and hence we can rewrite Eq. 8.33 as simply

$$\vec{F}_{\text{ext }1} = m\vec{a}_{\text{cm}} \tag{8.35}$$

or

$$\vec{a}_{\text{cm}} = \frac{\vec{F}_{\text{ext }1}}{m}. \tag{8.36}$$

Equation 8.36 is the equation of motion for the center of mass of the system. In words:

The center of mass of a two-object system accelerates as though both objects were located at the center of mass and the external force were exerted at that point.

You can now begin to see the power of the center-of-mass concept. Although we cannot determine the acceleration of the individual carts

without knowing more about the magnetic interaction, Eq. 8.36 allows us to calculate the motion of their center of mass *regardless* of the details of the magnetic interaction. As you will see in the next section, this is true not just for a system of two carts subject to a single external force, but even for a system of many interacting objects subject to many external forces, regardless of the interactions within the system and regardless of where on the system the external forces are exerted.

8.12 Systems of many interacting objects

With the analysis of a system comprising two objects under our belt, we are ready to look at the more complicated case of a system in which there are more than two objects. As in the preceding section, we begin by differentiating the momentum of the system with respect to time:

$$\frac{d\vec{p}}{dt} = \frac{d\vec{p}_1}{dt} + \frac{d\vec{p}_2}{dt} + \cdots = \Sigma\vec{F}_1 + \Sigma\vec{F}_2 + \cdots, \tag{8.37}$$

where $\vec{p}_1, \vec{p}_2, \cdots$ are the momenta of the individual objects in the system and $\Sigma\vec{F}_1, \Sigma\vec{F}_2, \cdots$ are the vector sums of the forces exerted on each of them. Each of these sums has two parts: one due to internal interactions and one due to external interactions. For object 1, for example, we have a sum of internal forces and a sum of external forces exerted on it:

$$\Sigma\vec{F}_1 = \Sigma\vec{F}_{\text{ext }1} + \Sigma\vec{F}_{\text{int }1} \tag{8.38}$$

and we can write similar expressions for all the other objects. Therefore the right side of Eq. 8.37 consists of a sum of external and internal forces:

$$\frac{d\vec{p}}{dt} = \Sigma(\vec{F}_{\text{ext }1} + \vec{F}_{\text{ext }2} + \cdots) + \Sigma(\vec{F}_{\text{int }1} + \vec{F}_{\text{int }2} + \cdots). \tag{8.39}$$

All internal forces drop out of Eq. 8.39, however. To see why, consider the case in which there are no external forces. Then the right side of Eq. 8.39 reduces to the sum of all internal forces:

$$\frac{d\vec{p}}{dt} = \Sigma(\vec{F}_{\text{int }1} + \vec{F}_{\text{int }2} + \cdots). \tag{8.40}$$

In this case, however, the system is isolated (because no external forces are exerted on it), and so we know that its momentum does not change. Therefore $d\vec{p}/dt = \vec{0}$ and Eq. 8.40 becomes

$$\vec{0} = \Sigma(\vec{F}_{\text{int }1} + \vec{F}_{\text{int }2} + \cdots). \tag{8.41}$$

In words, all the internal forces add up to zero! Therefore all of them drop out of the right side of Eq. 8.39, and the time rate of change of the momentum of the system is the sum of the external forces:

$$\frac{d\vec{p}}{dt} = \Sigma(\vec{F}_{\text{ext }1} + \vec{F}_{\text{ext }2} + \cdots) \equiv \Sigma\vec{F}_{\text{ext}}, \tag{8.42}$$

where $\Sigma\vec{F}_{\text{ext}}$ is the vector sum of all of the external forces exerted on the system. (See also the box "Cancellation of internal forces" on page 270.)

Cancellation of internal forces

There is another way to see that internal forces always add to zero. We consider here a system consisting of three particles, but the proof can be extended to any number of particles.

Each particle is subject to forces from the other two particles and to an external force:

$$\Sigma\vec{F}_1 = \Sigma\vec{F}_{\text{ext }1} + \vec{F}_{21} + \vec{F}_{31}$$
$$\Sigma\vec{F}_2 = \Sigma\vec{F}_{\text{ext }2} + \vec{F}_{12} + \vec{F}_{32}$$
$$\Sigma\vec{F}_3 = \Sigma\vec{F}_{\text{ext }3} + \vec{F}_{13} + \vec{F}_{23}.$$

Summing up all the forces, we get

$$\Sigma\vec{F}_1 + \Sigma\vec{F}_2 + \Sigma\vec{F}_3$$
$$= \Sigma\vec{F}_{\text{ext }1} + \Sigma\vec{F}_{\text{ext }2} + \Sigma\vec{F}_{\text{ext }3} + (\vec{F}_{12} + \vec{F}_{21})$$
$$+ (\vec{F}_{13} + \vec{F}_{31}) + (\vec{F}_{32} + \vec{F}_{23}). \qquad (1)$$

Remember, however, that two interacting particles exert on each other forces that are equal in magnitude but opposite in direction (Eq. 8.15). This means that the following equalities must hold:

$$\vec{F}_{12} = -\vec{F}_{21}$$
$$\vec{F}_{13} = -\vec{F}_{31}$$
$$\vec{F}_{32} = -\vec{F}_{23}.$$

So each term in parentheses in Eq. 1 vanishes, and only the external forces remain:

$$\Sigma\vec{F}_1 + \Sigma\vec{F}_2 + \Sigma\vec{F}_3$$
$$= \Sigma\vec{F}_{\text{ext }1} + \Sigma\vec{F}_{\text{ext }2} + \Sigma\vec{F}_{\text{ext }3}.$$

Using the definition of center-of-mass velocity again, we have $\vec{p} = m\vec{v}_{\text{cm}}$, where $m = m_1 + m_2 + \cdots$, so that

$$\frac{d\vec{p}}{dt} = \frac{d(m\vec{v}_{\text{cm}})}{dt} = m\frac{d\vec{v}_{\text{cm}}}{dt} \equiv m\vec{a}_{\text{cm}}. \qquad (8.43)$$

After substituting Eq. 8.43 into Eq. 8.42, we obtain the same result as for the two-object nonisolated system:

$$\Sigma\vec{F}_{\text{ext}} = m\vec{a}_{\text{cm}}. \qquad (8.44)$$

Now, however, this expression applies to many objects and many external forces.* What this result means is the following: If external forces are exerted on the objects in a many-object system, the system's center of mass moves as though all the objects in the system were concentrated at the center of mass and all external forces were exerted at that point.

The equation of motion for the center of mass of a nonisolated system is thus

$$\vec{a}_{\text{cm}} = \frac{\Sigma\vec{F}_{\text{ext}}}{m}. \qquad (8.45)$$

Note the surprising similarities between the equations for a single object (Eqs. 8.4 and 8.6),

$$\frac{d\vec{p}}{dt} = \Sigma\vec{F} \quad \text{and} \quad \Sigma\vec{F} = m\vec{a}, \qquad (8.46)$$

and those for the center of mass of a system of interacting objects (Eqs. 8.42 and 8.44),

$$\frac{d\vec{p}}{dt} = \Sigma\vec{F}_{\text{ext}} \quad \text{and} \quad \Sigma\vec{F}_{\text{ext}} = m\vec{a}_{\text{cm}}. \qquad (8.47)$$

*Equation 8.44 is less general than Eq. 8.42. The latter is always true, but the former holds only if the inertia m of the system is not changing. Equation 8.44 cannot be used, for instance, to calculate the acceleration of a rocket where the burning and high-speed ejection of fuel result in an appreciable change in the rocket's inertia.

The most significant difference between Eqs. 8.6 and 8.44 is that instead of the acceleration $\vec{a}$ of a single object, Eq. 8.44 contains the center-of-mass acceleration $\vec{a}_{cm}$ of the system. This subtle but very important difference is what makes it possible to apply this analysis to everyday objects. Any object can be regarded as a system of interacting particles; after all, objects are made up of many interacting atoms. Thus all the relationships we just derived apply to macro-sized objects.

One interesting application is to macro-sized objects that deform easily. For example, imagine pushing a down pillow (**Figure 8.23**). When you begin pushing on one edge of the pillow, your hand deforms the side it touches, meaning that part of the pillow is accelerating. While your hand is accelerating that side of the pillow, the other side of the pillow doesn't move at all. In other words, different parts of the pillow have different instantaneous velocities and different instantaneous accelerations. It therefore makes no sense, while deformation is taking place, to speak of *the* acceleration $\vec{a}$ of the pillow, and Eqs. 8.46 cannot be applied. On the other hand, we can always use Eqs. 8.47 to describe the motion of the pillow's center of mass. In other words, the momentum $\vec{p}$ of the pillow, the acceleration $\vec{a}_{cm}$ of its center of mass, and the vector sum of the forces exerted on it are still well-defined quantities.

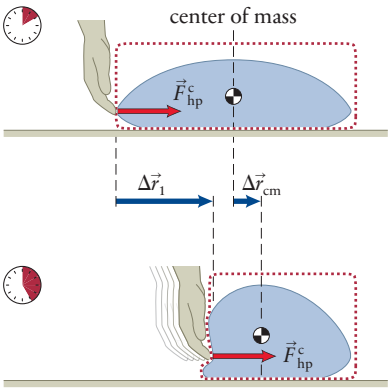
Perhaps the most remarkable aspect of this result is that we can always apply Eqs. 8.42 and 8.44 regardless of the internal interactions in a system. Even without any knowledge of these interactions, we can determine the motion of the center of mass of the pillow. What we give up by not considering the internal interactions is the ability to determine how the system deforms or what the motions of the different parts of the system are when the external force is exerted. In Figures 8.22 and 8.23, for example, we can determine only the displacement $\Delta\vec{r}_{cm}$ of the center of mass, not the displacement $\Delta\vec{r}_1$ of the point where the external force is exerted on the system, or the displacement of any other point.

The difference between Eqs. 8.46 and 8.47 vanishes when we consider a single object and ignore any deformation in it—that is to say, if we consider a **rigid object.** Although I argued in Section 8.6 that no object is truly rigid, most objects can be treated as being rigid for the purpose of describing their motion. For example, imagine pushing a stalled car toward the side of the road. For motion in one dimension, the front of the car moves the same way the back does—the minute compression of the car while your hands are pushing it is negligible. If you displace the back of the car by, say, 10 m, you can be pretty sure that the front of the car is also displaced by 10 m (give or take a fraction of the width of a hair). *All* parts of the car get displaced by 10 m in the same direction, and so the velocity and acceleration are the same for all parts of the car. For rigid objects we do not need to take into account any internal interactions or any difference in the displacement of various parts of the object.

Only for *deformable objects* (or for systems consisting of two or more rigid objects that can move relative to one another, like the two magnetically repelling carts in Figure 8.22) do you have to use the results derived in the last two sections.

8.20 (*a*) Show that Eqs. 8.47 reduce to Eqs. 8.46 when you consider a system of just one object. (*b*) Follow the procedure used to get from Eq. 8.21 to Eq. 8.24 for a system of many interacting objects.

Figure 8.23 As the external force exerted by the hand displaces a down pillow, the pillow deforms, and so different parts of the pillow are displaced by different amounts. For objects that deform, it makes no sense to speak of "the displacement" or "the acceleration" of the object. However, it does makes sense to speak of the displacement or the acceleration of the center of mass.



Chapter Summary

Characteristics of forces (Sections 8.1, 8.2, 8.3, 8.7)

Concepts When an object participates in one interaction only, the **force** exerted on the object is given by the time rate of change in the object's momentum.

A **contact force** is a force that one object exerts on another object only when the two objects are in physical contact.

A **field force** is a force (such as gravity) that one object exerts on another object without the requirement that the two objects be in physical contact.

When two objects interact, the forces they exert on each other form an **interaction pair,** and these forces have equal magnitudes but opposite directions.

Quantitative Tools The SI unit of force is the **newton** (N):

$$1\ \text{N} = 1\ \text{kg}\cdot\text{m/s}^2. \tag{8.9}$$

For an interaction pair of forces,

$$\vec{F}_{12} = -\vec{F}_{21}. \tag{8.15}$$

Some important forces (Sections 8.6, 8.8, 8.9)

Concepts A taut, flexible object (such as a spring, rope, or thread), when subjected to equal and opposite *tensile forces* applied at either end, experiences along its length a *stress* called **tension.** If the object is very light, the tension is the same everywhere in it.

Hooke's law relates the force that a spring exerts on a load to the distance the spring is stretched (or compressed) from its relaxed position.

Quantitative Tools The x component of the force of gravity exerted on an object of inertia m near Earth's surface is

$$F^G_{\text{E}o\,x} = -mg, \tag{8.17}$$

where the minus sign indicates that the force is directed downward.

Hooke's law: If a spring is stretched (or compressed) by a *small* distance $x - x_0$ from its unstretched length x_0, the x component of the force it exerts on the load is

$$(F_{\text{by spring on load}})_x = -k(x - x_0), \tag{8.20}$$

where k is called the **spring constant** of the spring.

Effects of force (Sections 8.1, 8.4, 8.5, 8.7)

Concepts The vector sum of the forces exerted on an object is equal to the time rate of change of the momentum of the object.

The **equation of motion** for an object relates the object's acceleration to the vector sum of the forces exerted on it.

Newton's laws of motion describe the effects forces have on the motion of objects.

If an object is at rest or moving with constant velocity, it is in **translational equilibrium.** In this case, the vector sum of the forces exerted on it is equal to zero.

A **free-body diagram** for an object is a sketch representing the object by a dot and showing all the forces exerted *on* it.

Quantitative Tools Vector sum of forces exerted on an object:

$$\Sigma\vec{F} \equiv \frac{d\vec{p}}{dt}. \tag{8.4}$$

Equation of motion:

$$\vec{a} = \frac{\Sigma\vec{F}}{m}. \tag{8.7}$$

If the inertia m of an object is constant, Newton's second law is usually written as

$$\Sigma\vec{F} = m\vec{a}. \tag{8.6}$$

For two interacting objects, 1 and 2, Newton's third law of motion states that

$$\vec{F}_{12} = -\vec{F}_{21}. \tag{8.15}$$

Impulse (Section 8.10)

Concepts The **impulse** $\vec{J}$ produced by forces exerted on an object is the product of the vector sum of the forces and the time interval over which the forces are exerted. The impulse delivered to the object is also equal to the change in its momentum.

Quantitative Tools For a constant force,

$$\Delta\vec{p} = \vec{J} = \left(\Sigma\vec{F}\right)\Delta t. \qquad (8.24, 8.25)$$

For a time-varying force,

$$\Delta\vec{p} = \vec{J} = \int_{t_i}^{t_f} \Sigma\vec{F}(t)\,dt. \qquad (8.26)$$

System of interacting objects (Sections 8.11, 8.12)

Concepts The center of mass of a system of objects accelerates as though all the objects were located at the center of mass and the external force were applied at that location.

Quantitative Tools Acceleration of a system of objects:

$$\vec{a}_{cm} = \frac{\Sigma\vec{F}_{ext}}{m}. \qquad (8.45)$$

Questions and Problems

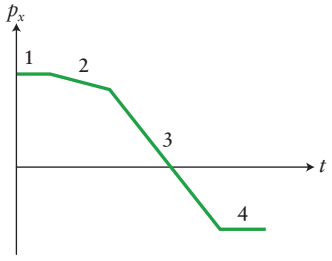
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

8.1 Momentum and force

1. A car merges onto a freeway, accelerating from 60 km/h to 80 km/h. As it does this, a bus passes by in the fast lane, travelling at a constant 100 km/h. Which vehicle has the greater magnitude of the vector sum of forces acting on it? •
2. You wish to move a crate. (*a*) Is there a minimum value for the magnitude of the pushing force you have to exert to set the crate in motion? (*b*) Is there a minimum value for the magnitude of the vector sum of forces that must be exerted on the crate to set it in motion? •
3. Sketch the $\Sigma F_x(t)$ curve for the vector sum of forces exerted on the object whose x component of momentum is plotted in Figure P8.3. Curve sections are numbered for reference. ••

Figure P8.3



4. The x component of the vector sum of forces exerted on a cart moving on a low-friction track is measured in the laboratory with the aid of a spring-loaded sensor and a monitoring computer. From the data in the table, sketch the $\Sigma F_x(t)$ curve of the cart and use your graph to estimate the change in the cart's momentum. ••

F (N)	t (s)
0.00	0.00
0.15	0.020
0.32	0.040
0.50	0.060
0.57	0.080
0.37	0.10
0.21	0.12
0.11	0.14
0.070	0.16
0.040	0.18

5. If a single known force is exerted on an object of unknown initial velocity for a certain time interval, you know the change in the object's momentum. However, you don't know its change in kinetic energy. Why not? •••
6. Consider a coal mining cart being loaded with coal as it moves along a horizontal track. Suppose that a tractor exerts a constant horizontal force on this cart. Is there any situation in which the cart-coal system would not accelerate? •••

8.2 The reciprocity of forces

7. A worker pushes a box in a factory. In each case, decide which force has the greater magnitude: the force exerted by the worker on the box or the force exerted by the box on the worker. (*a*) The box is heavy and does not move no matter how hard the worker pushes. (*b*) Some contents are removed, and now when pushed the box slides across the floor at constant speed. (*c*) The worker pushes harder, and the box accelerates. •

8. In many science fiction movies, someone is portrayed moving something with her or his mind—an amazing feat called *telekinesis*. What does your knowledge of forces predict about the effect of a telekinetic event on the gifted person? •
9. You jump into the air from a level floor. Which of the following are exerting a force on you: (*a*) the floor, (*b*) the Earth, (*c*) the air, (*d*) your feet? ••
10. During a tennis volley, a ball that arrives at a player at 40 m/s is struck by the racquet and returned at 40 m/s. The other player, realizing that the ball is out of bounds, catches it in her hand. Assuming the time interval of contact is the same in both cases, compare the force exerted by the first player's racquet on the ball with the force exerted by the second player's hand on the ball. ••

8.3 Identifying forces

11. Two soccer players kick two identical balls at the same time. Player 1 sends his ball at an impressive 90 km/hr at an angle of 30° above the horizontal. Player 2 kicks at 70 km/hr at 45°. Once in mid-air, which ball has a larger force acting on it? •
12. The force of friction is one of the most common forces that affects our everyday lives, and there are many situations where we either want to minimize or maximize it. Consider a children's playground and think of a situation where a large force of friction is important, and also another situation where a small force of friction is desirable. •
13. A boy bounces on a trampoline. When the boy is up in the air at the highest point, can you say that the total force on him is zero? ••
14. (*a*) Can an object ever exert both a contact force and a field force on another object? (*b*) Can an object ever exert a contact force on something without exerting a field force on that same object? ••

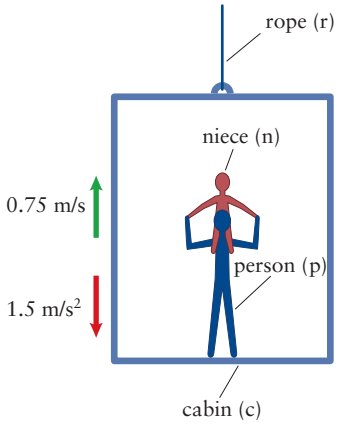
8.4 Translational equilibrium

15. You and two other friends are pushing a car over the rough road ahead. A physicist observing the car's motion claims there is no net force acting on the car. Explain. •
16. You are discussing physics homework with a friend while the two of you are riding your bikes. Distracted, you collide with a parked car. The car doesn't move. Your friend says, "Of course not! The force of the collision wasn't enough to overcome the inertia of the car." How do you respond? •
17. What interaction does a bathroom (spring) scale actually measure? ••
18. A table sits on the floor, and a vase containing water and flowers sits on top of the table. All objects are at rest. Identify all forces exerted on the table, and describe how these forces are related. ••

8.5 Free-body diagrams

19. You want to climb up a rope attached to a tree branch overhead. You test the strength of the rope by holding it while slowly lifting your feet off the ground; the rope holds your weight. However, as you start to climb rapidly up the rope, it suddenly breaks. Your weight has not changed, so why did the rope break? •
20. When you are standing motionless on the ground, your feet exert a force on Earth. Why doesn't Earth move away from you? ••
21. You push on a crate, and it starts to move—but you don't. Draw a free-body diagram for you and one for the crate; show only the horizontal forces. Then use the diagrams and your knowledge about forces and interactions to explain why the crate moves but you don't. ••
22. A ball falling near the surface of the Earth where air friction is not negligible eventually slows down to a constant "terminal" velocity. How is that possible given that frictional force increases with increasing velocity? ••
23. You are carrying your 3-year-old niece on your shoulders in an elevator cabin that is moving upward with velocity 0.75 m/s and constant downward acceleration of 1.5 m/s^2 (Figure P8.23*). Your inertia is 60 kg, your niece's inertia is 25 kg, and the cabin's inertia is 95 kg. (*a*) Draw free-body diagrams for the elevator cabin, your niece, and yourself, labeling the forces using the notation introduced in section 8.5. Make sure the lengths of the force vectors are appropriate to the magnitudes of the forces. (*b*) State the magnitudes of as many forces as possible, and identify all the forces that form interaction pairs. Explain briefly how you obtained these magnitudes. •••

Figure P8.23



8.6 Springs and tension

24. A delivery person in an elevator is holding a package by an elastic cord. (Don't ask why.) (*a*) What happens to the length of the cord when the elevator accelerates upward? Draw the free-body diagram for the package in this case. (*b*) What happens to the cord length when the elevator slows to a stop after its ascent? Draw the free-body diagram for the package in this case. •
25. Suppose a rope that is hanging by one end from the ceiling has a large inertia. Is the tension in the rope uniform throughout the length? ••
26. Walking beside a pasture, you and a fellow student see a farmer pulling a mule with a rope and getting nowhere. Your friend says, "The force with which the mule is

pulling on the rope has the same magnitude as the force with which the farmer is pulling on the rope, but the two forces point in opposite directions. Because the two forces cancel, the tension in the rope is zero." How do you respond? ●●

27. In a game of tug of war, the blue team is beating the red team, and the red team is accelerating toward the mud pit between the two teams. Which has the greater magnitude, if either: the force the blue team exerts on the rope or the force the red team exerts on the rope? Assume that the gravitational force exerted on the rope is negligible compared with the forces exerted by the two teams on the rope. ●●

28. The standard coupling between railroad cars must be capable of withstanding the maximum tensile force exerted on any coupling in a given train. (*a*) If a locomotive is pulling ten cars and speeding up, on which coupling is the greatest tensile force exerted? (*b*) Is this tensile force due to stretching or compression? (*c*) If the locomotive is slowing the train down, on which coupling is the greatest tensile force exerted? (*d*) Is this tensile force due to stretching or compression? ●●

8.7 Equation of motion

29. A proton (inertia 1.67×10^{-27} kg) and an electron (inertia 9.11×10^{-31} kg) are being separately subjected to an electric force of 3.2×10^{-14} N. Compare their ensuing accelerations. ●

30. An object moving along the x-axis experiences a constant force in its direction of motion, $F_x = 56$ N. At $t = 1.2$ s, the object is moving at 2.0 m/s and at $t = 3.5$ s it is moving at 2.8 m/s. What is the inertia of the object? ●

31. Using this chapter, is it possible to figure out an object's past motion if you know its present position and velocity and the history of all the forces exerted on it? ●●

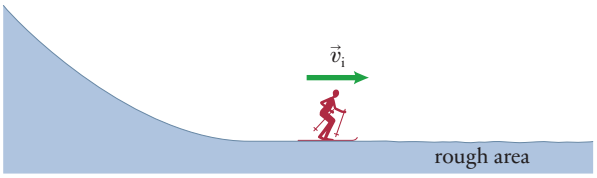
32. The force exerted by expanding gases is what propels a shell out of a gun barrel. Suppose a force that has an average magnitude F propels the shell so that it has a speed v at the instant it reaches the end of the barrel. What magnitude force is required to double the speed of the shell in the same gun? ●●

33. A 50-kg pneumatic hammer is fitted with a shock absorber at the top to lessen the impact on the worker. The vertical displacement of the hammer as a function of time is given by $x(t) = at^2 - bt^3$ (during the hammer's bounce), where $a = 15$ m/s^2 and $b = 20$ m/s^3. The positive x axis is directed upward. (*a*) Write an expression for the x component of the vector sum of forces exerted on the hammer as a function of time. (*b*) At which values of t is the x component of the vector sum of forces on the hammer positive, at which values of t is it negative, and at which values of t is it zero? ●●

34. You push a 30 kg shopping cart initially at rest (loaded with groceries) with a constant 12-N force. How far does it travel after pushing it for 4.5 s? ●●

35. In Figure P8.35, a 55-kg skier heads down a slope, reaching a speed of 27 km/h. They then slide across a horizontal snow field but hit a rough area. Assume the snow before the rough area is so slippery that you can ignore any friction between the skier and the snow. However the rough area exerts a significant constant friction force. If the skier comes to rest after travelling 50 m across the rough area, what is the magnitude of the friction force? ●●

Figure P8.35



36. A 2.34-kg cart on a long, level, low-friction track is heading for a small electric fan at 0.23 m/s. The fan, which was initially off, is turned on. As the fan speeds up, the magnitude of the force it exerts on the cart is given by at^2, where $a = 0.0200$ N/s^2. What is the speed of the cart 3.5 s after the fan is turned on? After how many seconds is the cart's velocity zero? ●●●

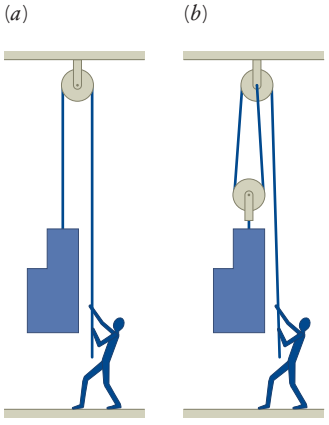
8.8 Force of gravity

37. Two blocks of the same inertia (10 kg) but different sizes are falling freely. Compare the force exerted by the Earth on the two blocks. ●

38. Estimate the magnitude of the force of gravity exerted on (*a*) a packet of sugar, (*b*) a cup of water, (*c*) a packed carry-on suitcase, (*d*) an adult gorilla, and (*e*) a midsize car. ●

39. The magnitude of the gravitational force exerted on the piano in Figure P8.39 is 1500 N. With how much force do you have to pull on the rope in order to hold up the piano (*a*) with the single pulley and (*b*) with the double pulley? ●

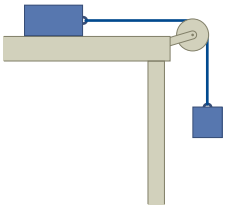
Figure P8.39



40. The piano in Figure P8.39 is being lowered with an acceleration of magnitude exactly $g/8$. To maintain this acceleration, with what force magnitude must the person pull on the rope (*a*) with the single pulley and (*b*) with the double pulley? ●●

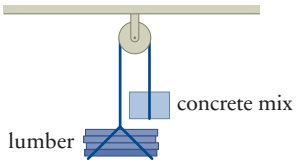
41. If the force of gravity exerted by Earth on a fully suited astronaut at Earth's surface is 1960 N, what is the magnitude of the gravitational force exerted on her (*a*) on Jupiter, where the acceleration due to gravity is 25.9 m/s^2, and (*b*) on the Moon, where the acceleration due to gravity is 1.6 m/s^2? ●●
42. You are collecting water from a well by lowering a bucket on a rope. The inertia of the empty bucket is 0.72 kg, and the inertia of the rope is negligible. (*a*) If you lower the empty bucket with a downward acceleration of 4.8 m/s^2, what is the tension in the rope? (*b*) You now pull the filled bucket, with total inertia 6.6 kg, back to the top of the well. If the maximum tension that the rope can support is 100 N, what is the maximum upward acceleration for the bucket? ●●
43. Figure P8.43 shows two blocks, one of which is placed on a low-friction table, and supports, by means of a massless rope that runs over a low-friction pulley, the other block, which is suspended in mid-air. If you hold the hanging block, and then release it, how big an inertia should this block have for the system to slide? What would the acceleration of the system be, in terms of *g*, *m*, and M? ●●

Figure P8.43



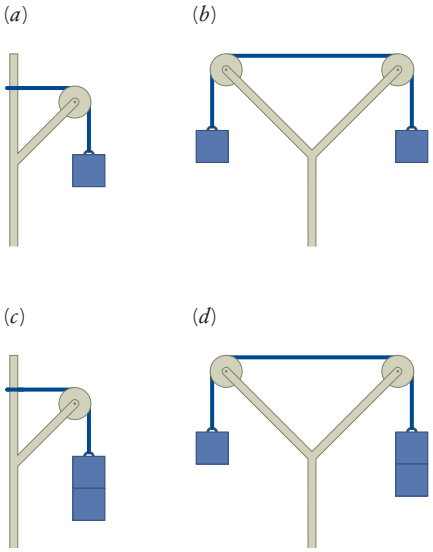
44. Assume that the block on the table in Figure P8.43 has twice the inertia of the hanging block. (*a*) You give the block on the table a push to the right so that it starts to move. If the magnitude of the frictional force exerted by the table on the block is half the magnitude of the gravitational force exerted on the hanging block, what is the acceleration of the hanging block after you have stopped pushing the block on the table? (*b*) What would this hanging-block acceleration be if you had pushed the block on the table to the left instead? Consider only a short time interval after you stop pushing. ●●
45. Consider figure P8.43. Let the inertia of the block on the table be M and that of the hanging bock be M/4. If the coefficient of kinetic friction between the table and the block is 0.2, calculate the tension in the rope and the acceleration of the hanging block. ●●
46. A bag of concrete mix is tied to one end of a rope that is looped over a pulley attached to the ceiling, and the other end of the rope is tied to a pallet of lumber (Figure P8.46). The concrete mix has a greater inertia than the lumber. When both loads are released from rest so that the system is free to move, compare the tension in the rope with the magnitudes of the gravitational forces exerted by Earth on the lumber and on the concrete mix. Ignore any friction in the pulley. ●●

Figure P8.46



47. All blocks in Figure P8.47 are identical and you can ignore any friction in the pulleys. Rank the configurations in increasing order of tension in the rope. (Hint: Use free-body diagrams.) ●●

Figure P8.47



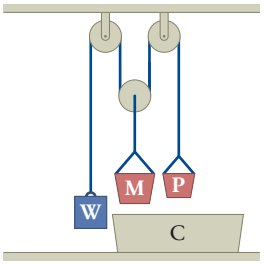
48. Picture an object in free fall. If the leading face of the object (which means the face closest to the ground) has a large surface area, air resistance becomes important. For low speeds, the force due to air resistance may be approximated as proportional to the object's velocity according to the expression $\vec{F}^{a} = -b\vec{v}$, where *b* is a number that depends on the shape and volume of the object and on the density of the air. (*a*) Show that, for an object of inertia *m*, there is a limiting speed (called the *terminal speed*) given by $v_t = mg/b$. (*b*) Show that, if the object is released from rest, its speed as a function of time is given by

$$v(t) = v_t[1 - e^{-(b/m)t}].$$

(*c*) What are the units of m/b? Sketch a graph of *v* as a function of time, with the horizontal axis in units of multiples of m/b. ●●●
49. At a munitions factory, chemicals M and P have to be added to a vat containing chemical C. Chemicals M and

P have to be added simultaneously to prevent an accidental detonation. Your task, as the chemical engineer in charge, is to come up with a way of doing this, and you devise the pulley scheme shown in Figure P8.49, where block W is a counterweight whose inertia controls the speed at which the buckets containing chemicals M and P move. The idea is that when a strap holding block W at rest is released, the block rises and the buckets containing M and P descend into the vat. In order to have the two buckets have an identical acceleration of magnitude $g/3$, what must be the ratios of the inertias m_W, m_M, and m_P? Ignore any friction in the pulleys. •••

Figure P8.49



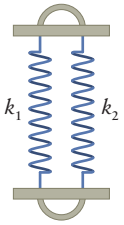
8.9 Hooke's law

50. A 66-kg person experiences a gravitational force of about 660 N. Yet, if this person were to jump onto a spring scale, the scale would briefly read about 2400 N. Why does the scale register a force so much larger than the gravitational force exerted on the person? •

51. A 27-kg child stands in the center of a trampoline. (*a*) If the trampoline center is 0.32 m lower than before they got on, what is the spring constant of the trampoline? (*b*) Assuming the trampoline acts like a spring, how low would their 65-kg father sink if he stood alone in the center? •

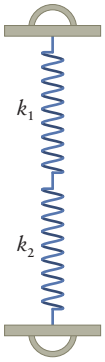
52. Two springs have spring constants k_1 and $k_2 > k_1$. Connected as shown in Figure P8.52, they act like one spring. Compute the spring constant of the combination, and determine whether it is smaller than k_1, equal to k_1, between k_1 and k_2, equal to k_2, or greater than k_2. ••

Figure P8.52



53. Two springs have spring constants k_1 and $k_2 > k_1$. Connected as shown in Figure P8.53, they act like one spring. Compute the spring constant of the combination. Is the combination smaller than k_1, equal to k_1, between k_1 and k_2, equal to k_2, or greater than k_2? ••

Figure P8.53



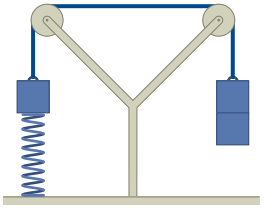
54. Two springs are hooked together end to end. When a 4.0-kg brick is suspended from one end of the combination, the combination stretches 0.15 m beyond its relaxed length. (*a*) What is the spring constant of the combination? (*b*) If the top spring stretches 0.10 m, what is the spring constant of each spring? ••

55. When a 5.0-kg box is hung from a spring, the spring stretches to 50 mm beyond its relaxed length. In an elevator accelerating upward at 2.0 m/s^2, how far does the spring stretch with the same box attached? ••

56. Suppose that when you lie on your bed, your body is supported by 30 of the springs in the mattress. Assuming that all the springs are identical and that your body presses on each spring equally, estimate their spring constant. (Experiment: Lie on your bed and see how far down you sink.) ••

57. In the modified Atwood machine shown in Figure P8.57, each of the three blocks has the same inertia m. One end of the vertical spring, which has spring constant k, is attached to the single block, and the other end of the spring is fixed to the floor. The positions of the blocks are adjusted until the spring is at its relaxed length. The blocks are then released from rest. What is the acceleration of the two blocks on the right after they have dropped a distance d? Ignore any friction in the pulleys. ••

Figure P8.57

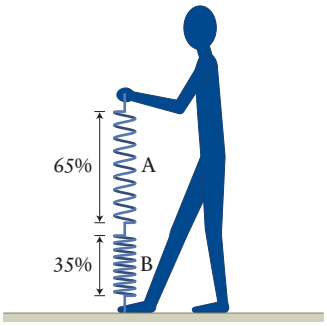


58. Your boss hands you a bag containing four springs, A, B, C, and D, and tells you that they all have the same relaxed length. He wants you to rank them by the values of their spring constants. He locks you in a room with only the springs, a ruler, and a pad and pencil, and tells you to knock when you're done. Each spring has a hook on one end and a handgrip on the other. There is nothing in the room you can use to measure inertia.

After thinking for a few minutes, you hook springs A and B together, stand on one handgrip with your foot, and pull on the other handgrip with your hand until the

combination spring has been stretched to double its original length (Figure P8.58). You notice that spring A makes up 65% of the stretched length and spring B makes up the other 35%.

Figure P8.58



You work your way through the rest of the springs and tabulate the results:

Combination	Percent of stretched length
A/B	A 65, B 35
A/C	A 70, C 30
A/D	A 67, D 33

Based on these results, what are the numerical values in the ratio k_A: k_B: k_C: k_D? •••

59. A 5.0-kg red book is suspended from a spring attached to the ceiling of an elevator that is accelerating downward. Attaching the book causes the spring to stretch 71 mm beyond its relaxed length. You also have a yellow book of unknown inertia. When you attach the yellow book and the red one to the spring at the same time, the spring stretches 110 mm beyond its relaxed length. When the elevator is at rest, the combination of the two books stretches the spring 140 mm beyond its relaxed length. (*a*) What is the spring constant? (*b*) What is the inertia of the yellow book? (*c*) What is the elevator's acceleration? •••

8.10 Impulse

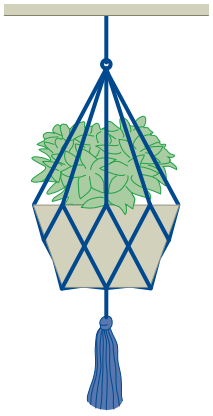
60. Use the relationship between force and impulse to explain how padded boxing gloves protect a boxer's hands. •

61. If you try to pound a nail into a board with a 2.5-kg rubber mallet, you will have less luck than if you use a 0.8-kg steel hammer. Why? •

62. In a baseball game, a batter hits the 0.150-kg ball straight back at the pitcher at 180 km/h. If the ball is traveling at 160 km/h just before it reaches the bat, what is the average force exerted by the bat on it if the collision lasts 5.0 ms? •

63. In Figure P8.63, a heavy potted plant hangs by an enveloping hanger made of heavy yarn. (*a*) If you pull slowly on the tassel at the bottom, where is the hanger most likely to break? Why? (*b*) If you yank quickly on the tassel, where is the hanger most likely to break? ••

Figure P8.63



64. Estimate the impulse delivered by a hammer to a nail being driven into a wood plank. ••

65. An 1100-kg car stops at a red light, going from 85 km/h to stationary in 20 s. (*a*) What is the impulse delivered to the car? (*b*) What is the average vector sum of forces exerted on the car? ••

66. A 2-kg block resting on a low-friction floor is subjected to the time-dependent force $F(t) = 0.5t^2$. The block starts out 1.0 m away from the origin. (*a*) What is speed of the object at $t = 2.0$ s? (*b*) What is the position of the object at $t = 2.0$ s? ••

67. Just before hitting the ground, a partially inflated 0.625-kg basketball has a speed of 3.30 m/s. Then it loses half of its kinetic energy as it bounces. (*a*) What is the ball's speed immediately after it bounces? (*b*) If the ball is in contact with the ground for 9.25 ms, what is the magnitude of the average force exerted by the ground on the ball? ••

68. A 2.3-kg ball dropped from a height of 2.0 m onto a steel plate rigidly attached to the ground bounces back to a height of 1.7 m. (*a*) What is the impulse delivered to the ball by the plate? (*b*) What is the coefficient of restitution of the collision? ••

69. A 0.050-kg egg falls off a table that is 1.0 m high and lands on a tile floor. (*a*) What is the impulse delivered to the egg by the floor? (*b*) The egg lands on its side, and at the instant of impact the part of the shell farthest from the floor is 40 mm above the floor. The collision continues until the center of mass is on the floor. Estimate the time interval over which the collision lasts. (*c*) What is the average force exerted by the floor on the egg? (*d*) If the floor were carpeted, the collision time interval might be extended by as much as a factor of four as the carpet "gives" under the egg. What would the average force exerted on the egg be then? ••

70. When it crashes into a bridge support that does not move, a car goes from 80 km/h to 0 in 1.23 m. (*a*) What is the impulse delivered to the 70-kg driver by the seat belt, assuming the belt makes the driver's motion identical to the car's motion? (*b*) What is the average force exerted by the belt on the driver? (*c*) If the driver were not wearing the seat belt and flew forward until he hit the steering wheel, which stopped him in 0.0145 s, what would be the average force exerted by the steering wheel on him? (*d*) Would he be likely to survive the crash described in part *c*? Justify your answer by comparing

the force exerted by the steering wheel on the driver to the gravitational force exerted on him. •••

71. A volleyball player serves the ball. The 0.27-kg ball was moving straight up and is at the high point of its trajectory when she hits it in a purely horizontal direction. The magnitude of the force exerted on the ball while her hand is in contact with it is given by

$$F(t)_{\text{hb}}^{\text{c}} = at - bt^2,$$

where $a = 3.0 \times 10^5$ N/s and $b = 1.0 \times 10^8$ N/s^2. Her hand is in contact with the ball for 3.0 ms. What is the magnitude of (*a*) the impulse delivered to the ball, (*b*) the average force exerted by her hand on the ball, and (*c*) the maximum force exerted by her hand on the ball? (*d*) What is the velocity of the ball immediately after her hand loses contact with it? •••

8.11 Systems of two interacting objects

72. A 70-kg student is falling toward Earth. (*a*) Draw one free-body diagram for Earth and one for the student. Assume that air resistance can be ignored. (*b*) Calculate the student's acceleration and the vector sum of forces exerted on him. (*c*) Earth's inertia is 5.97×10^{24} kg. What is the magnitude of the planet's acceleration toward the student? •

73. Box A, which has an inertia of 2 kg, and box B, which has an inertia of 3 kg, are 2.0 m apart on a low-friction floor. Box B is pushed toward box A by a steady force of 10.0 N. What is the acceleration of (*a*) box A and (*b*) the center of mass of the system before they collide? •

74. Two 0.500-kg carts are 100 mm apart on a low-friction track. You push one of the carts with a constant force of 2.00 N directed so that the cart you push moves away from the other cart. Determine the acceleration of the center of mass of the two-cart system when (*a*) the carts do not interact and (*b*) the carts are connected by a spring of spring constant $k = 300$ N/m. ••

75. A force of magnitude F is exerted on the leftmost face of two blocks sitting next to each other on a slippery surface, with two of their faces touching. The inertia of the block on the right is m, and the inertia of the block on the left is $2m$. (*a*) What is the magnitude of the acceleration of each block? (*b*) What are the magnitude and direction of the contact force the left block exerts on the right block? (*c*) What are the magnitude and direction of the contact force the right block exerts on the left block? (*d*) How do your answers to parts *a*, *b*, and *c* change if the positions of the two blocks are interchanged? ••

76. A 2.0-kg cart and an 8.0-kg cart are connected by a relaxed, horizontal spring of spring constant 300 N/m. You pull the 8.0-kg cart with some constant horizontal force. The separation between the carts increases for a short time interval, then remains constant as you continue to pull and the spring is stretched by 0.100 m. (*a*) What pulling force did you exert? (*b*) If you instead pull the 2.0-kg cart, what force must you exert to get the same stretch in the spring? ••

77. A 1500-kg truck and a 1000-kg car are parked with their rear bumpers nearly touching each other in a level parking lot. Both vehicles have their brakes off so that they are free to roll. A man sitting on the rear bumper of the truck exerts a constant horizontal force on the rear bumper of the car with his feet, and the car accelerates at 1.2 m/s^2. (*a*) What are the magnitude and direction of the acceleration of the center of mass of the car-truck system? (*b*) What vector sum of forces (magnitude and direction) is exerted on each vehicle? (*c*) What are the magnitude and direction of the acceleration of the truck? Ignore any friction between the tires and the parking lot surface. ••

78. A 70-kg skier is being towed on a rope behind a 450-kg snowmobile on a smooth, snow-covered surface at 10 m/s when the snowmobile hits a patch of muddy ground that brings it to a halt in 10 m. (*a*) What is the average acceleration of the snowmobile while it is slowing? (*b*) What is the average acceleration of the center of mass of the skier-snowmobile system? ••

79. A 4.0-kg block and a 2.0-kg block are connected to opposite ends of a relaxed spring of spring constant 300 N/m. The blocks are pushed toward each other for 0.50 s. The 4.0-kg block is pushed with a force of 50 N, and the 2.0-kg block is pushed with a force of 20 N. Ignore the inertia of the spring. (*a*) What is the acceleration of the center of mass of the two-block system while this pushing is going on? (*b*) When the blocks are released, what is the velocity of the 2.0-kg block if the 4.0-kg block is moving away from it at 5.0 m/s in the Earth reference frame? ••

8.12 Systems of many interacting objects

80. You push your 0.70-kg pillow across your bed with a constant force of 10 N. The bed provides a frictional force of 6.0 N. What is the acceleration of the center of mass of the pillow? •

81. You are holding a basketball while standing on a bathroom scale. You now throw the basketball directly upwards, and catch it on its way down. Describe how the reading on the scale changes from its initial value (*a*) as you are throwing the basketball, (*b*) after the basketball has left your hands, (*c*) while you are catching the basketball again. ••

82. You build a 2.0-kg mound of mashed potatoes on a kitchen scale (which is a spring scale). You then use a large spoon to press on the top of the mound at a constant force of 3.0 N. While you are doing this, the scale reading increases by 5.0%. (*a*) What is the acceleration of the center of mass of the potatoes? (*b*) Describe what happens to the shape of the mound and the scale reading as time goes by. ••

83. You use a rope of length ℓ to tie the front of a 1500-kg truck to the rear of a 1000-kg car and then use an identical rope of length ℓ to tie the front of the car to the rear of a 500-kg trailer. Each rope will break if its tension exceeds 2000 N. You attach a strong cable to the rear of the truck and use a winch to pull it with a force of magnitude F. (*a*) What is the acceleration of the three-vehicle system? (*b*) What is the maximum pulling force you can exert on the truck before a rope breaks? (*c*) If you now attach the cable to the front of the trailer and again pull the system, what is the maximum force with which you can pull it without snapping a rope? ••

84. A tugboat pulls two barges down the river. The barge connected to the tugboat, carrying coal, has an inertia of 2.0×10^5 kg, and the other barge, carrying pig iron, has an inertia of 3.0×10^5 kg. The resistive force between the coal barge and the water is 8.0×10^3 N, and the resistive force between the pig iron barge and the water is 10×10^3 N. The common acceleration of all three boats is 0.40 m/s^2. Even though the ropes are huge, the gravitational force exerted on them is much smaller than the pulling forces. (*a*) What is the tension in the rope that connects the tugboat to the coal barge? (*b*) What is the tension in the rope that connects the two barges? (*c*) Repeat parts *a* and *b* for the case in which the order of the barges is reversed. ••

85. A farmer is trying to haul three 500-kg trailers connected together with identical ropes that will break if their tension exceeds 2000 N. The first trailer is connected to his tractor with a rigid hitch. (*a*) If the tractor has to overcome a frictional force of 900 N per trailer to start them moving, what is the minimum force the tractor must exert on the first trailer to get them all rolling if the ropes are taut? (*b*) What is the maximum force the tractor can exert without breaking any rope? ••

86. A train is comprised of an engine car of inertia M, and 3 more cars of inertias m_1, m_2, and m_3, respectively. The acceleration of the train is *a*. Friction can be ignored. ••
   (*a*) What is the net force acting on the train?
   (*b*) What is the force in the coupler pulling the last car (m_3)?
   (c) What is the force in the coupler between the engine and the first car (m_1)?

87. A red 10-kg cart is connected to a 20-kg cart by a relaxed spring of spring constant 60 N/m. The 20-kg cart is resting against another 10-kg cart, this one blue. All are on a low-friction track. You push the red 10-kg cart to the right, in the direction of the 20-kg cart, with a constant force of 10 N. (*a*) What is the acceleration of the center of mass of the three-cart system? (*b*) What is the acceleration of each cart the instant you begin to push? (*c*) What is the vector sum of the forces on each cart when the spring is compressed 100 mm? •••

Additional Problems

88. A toy wagon initially at rest is pulled by a child from one end of a driveway to the other end. The magnitude of the force the wagon exerts on the child is the same as the magnitude of the force the child exerts on the wagon. So, how does the child get the wagon moving? •

89. In years past, car frames were built as stiff as possible, but today's cars are built with "crumple zones" in the front and rear portions of the frame. What is the purpose of this design change? •

90. An 80-kg man rides in an elevator from the ground to an upper floor. The elevator goes from rest to 10 m/s in 5 seconds, and then continues at the same constant speed for 10 seconds, before slowing down to a stop in 5 seconds. How much push from the elevator floor does the man feel during (*a*) the initial acceleration, (*b*) the constant-speed phase, and (*c*) the final deceleration? ••

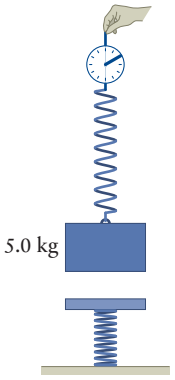
91. You want to hang an object from the ceiling of an elevator that has a maximum acceleration of 4.0 m/s^2. (*a*) If you hang the object with fishing line that supports 45 N of force, what is the maximum inertia the object can have if the line is not to break? (*b*) What combination of slowing down, speeding up, going up, and going down of the elevator causes the greatest force to be exerted on the fishing line? ••

92. You need to lower a 45-kg safe from a window to the bed of a 1800-kg truck. You have a rope that is just long enough, but it will support no more than 42 kg. What must you be careful to do (or not to do) so that the rope does not break? ••

93. A 10-kg cart is connected to a 20-kg cart by a relaxed spring of spring constant 1000 N/m, and both carts are placed on a low-friction track. You push the 10-kg cart in the direction of the 20-kg cart with a constant force of 10 N. (*a*) What is the acceleration of the center of mass of the two-cart system at any instant? (*b*) What is the acceleration of each cart at the instant you begin to push? (*c*) What is the acceleration of each cart when the spring has its maximum compression? ••

94. A 5.0-kg block suspended from a spring scale is slowly lowered onto a vertical spring (Figure P8.94). (*a*) What does the scale read before the block touches the vertical spring? (*b*) If the scale reads 40 N when the bottom spring is compressed 30 mm, what is *k* for the bottom spring? (*c*) How far does the block compress the bottom spring when the scale reads 0? ••

Figure P8.94



95. You need to lift a heavy load of inertia m_ℓ upward with an acceleration of magnitude $g/8$. You have two pulleys of inertia $m_P \ll m_\ell$, one attached to a beam above the loading dock and another that can be attached to the load. You consider a block-and-tackle system, which has two pulleys, but you realize that these are double pulleys (each has two wheels side by side on a common axle) and you have a long piece of rope. ••• CR

96. You and your uncle are fishing. You are using 22-N fishing line, which means the line can support that much tension without breaking. Your uncle challenges you to catch the biggest fish you can land in the boat with that line, with the provision that you have to hoist the fish

out of the water with the line and not with a net. You finally hook a fish that looks to have the right inertia, and you start to reel it in slowly. Just as you get the fish up into the air, though, the line snaps and the fish gets away. You accuse your uncle of having behind your back switched your pole to a line weaker than the 22-N line you thought you were working with, but he denies it. He then points out that you were accelerating the fish upward at 1.0 m/s^2. ●●● CR

97. Your 70-kg friend always thinks he can dominate you in contests of strength because, even though you are a very athletic gymnast, your inertia is 63 kg. This time, while the gym class watches, he has attached a pulley to the ceiling of the gym, passed a rope over it, and tied one end of the rope around his waist. He hands you the other end of the rope, which is not quite long enough to reach the floor, and bets you that you cannot lift him off the ground by pulling on the rope. You realize that gravitational forces will favor your friend if the two of you are hanging on opposite sides of the pulley, but then you recall your amazing arm strength gained from years of work on gymnastic apparatus. ●●● CR

98. A burst of compressed air pushes a pellet out of a blowpipe. The force exerted by the air on the pellet is given by $F(t) = F_0 e^{-t/\tau}$, where τ is called a *time constant* because it has units of time. (*a*) What does F_0 represent? (*b*) What is the momentum of the pellet after an interval equal to one time constant has elapsed? (*c*) What is the momentum after an interval equal to five time constants has elapsed? (*d*) What is the final momentum of the pellet after a long time interval has elapsed? (*e*) If $\tau = 0.50$ ms, after how long a time interval has the momentum reached 95% of its final value? ●●●

99. *Thrust* is the force that pushes a rocket forward. It is a force exerted on the rocket as the engine expels hot gases from the rear of the rocket. For most rockets, thrust is variable rather than constant during the "burn time" (the time interval during which fuel is consumed and expelled). However, the thrust of a certain rocket is nearly constant during part of its burn time, beginning when the rocket has consumed 40% of its fuel and continuing until the rocket has consumed about 90% of its fuel. As the rocket consumes and expels its fuel, its inertia decreases. Describe the rocket's acceleration during this interval of constant thrust. ●●●



9

Work

CONCEPTS

9.1 Force displacement
9.2 Positive and negative work
9.3 Energy diagrams
9.4 Choice of system

QUANTITATIVE TOOLS

9.5 Work done on a single particle
9.6 Work done on a many-particle system
9.7 Variable and distributed forces
9.8 Power

As you learned in Chapter 8, a system subject to external forces—that is to say, forces exerted by objects outside the system on objects inside the system—is not isolated. The relationship between the momentum of a nonisolated system and the vector sum of the forces exerted on that system is simple: The vector sum of the forces is equal to the time rate of change in the system's momentum (Eq. 8.4). In addition to affecting a system's momentum, external forces can change the energy of the system. In the simplest case, a single force exerted on a system consisting of a single object causes that object to accelerate, and so the force changes the object's kinetic energy. Forces can also change the physical state of objects (think of compressing or stretching a spring or crumpling a piece of paper by exerting a force on it). Because the physical state of an object is related to the object's internal energy, we see that forces can change not only an object's kinetic energy but its internal energy as well.

To describe these changes in energy, physicists use the concept of **work:**

Work is the change in the energy of a system due to external forces.

When an external force $\vec{F}_{\text{ext}}$ causes a change in the energy of a system, physicists refer to this change either as "the *work done by* the force $\vec{F}_{\text{ext}}$ on the system" or as "the *work done by* the object or person exerting force $\vec{F}_{\text{ext}}$ on the system."

Because work is a change in energy, the unit for work is the same as that for energy—the joule.

Note that this definition of work is much more limited than the meaning of the word *work* in ordinary language, which refers to any physical or mental effort or activity. You may be "doing work" for your physics class while reading this book, but there is no work involved in the physics sense.

9.1 Force displacement

Work amounts to a mechanical transfer of energy either from a system to its environment or from the environment to the system. For instance, if you (part of the "environment") push a friend on a bicycle (system is rider plus bicycle) and your friend accelerates, the force exerted by you on your friend transfers energy from your body to hers, increasing her kinetic energy. If you compress a spring, you transfer energy from your body to the spring, where it is stored in the form of elastic potential energy. In each case, the change in the system's energy is caused by an external force—a force exerted by you on the system.

If you wonder why a force has to be external to do work on a system—that is, to change the energy of the system—remember that forces represent interactions between objects. Any interaction between objects within the system rearranges the energy between these objects but doesn't change the overall amount of energy in the system. Only an interaction between an object inside the system and an object outside it—an external interaction—can transfer energy across the system boundary.

Do external forces *always* cause a change in the energy of a system? To answer this question, it is helpful to consider an example.

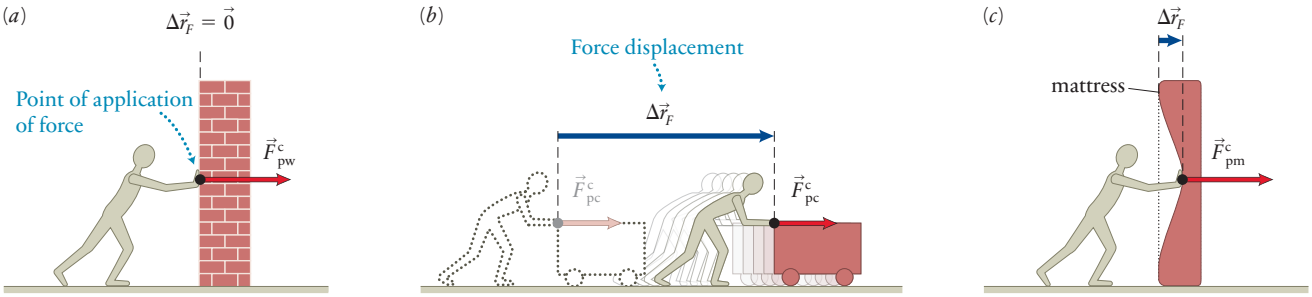
9.1 Imagine pushing against a brick wall as shown in Figure 9.1*a*. (Assume the wall is undeformable—a fair assumption at macroscopic scales.) (*a*) Considering the wall as the system, is the force you exert on it internal or external? (*b*) Does this force accelerate the wall? Change its shape? Raise its temperature? (*c*) Does the energy of the wall change as a result of the force you exert on it? (*d*) Does the force you exert on the wall do work on the wall?

As this checkpoint illustrates, external forces exerted on a system do not always do work on the system. The other two examples in **Figure 9.1**, however, do involve changes in energy, and hence in these cases work is done by your external force on the system. If, as shown in Figure 9.1*b*, you push against something that has wheels and so can accelerate it, its kinetic energy changes. If you push against a deformable object, the object changes shape and so its potential energy changes (Figure 9.1*c*).

Why is the work zero in Figure 9.1*a* but nonzero in parts *b* and *c*? The key to answering this question lies in looking at the *point of application* of each force (the point at which the force is exerted on the system, as indicated

CONCEPTS

Figure 9.1 An external force does work on a system only if it causes the point at which it is applied (represented by a black dot) to move.



Point of application does not move, so force does no work on wall.

Point of application moves, so force does work on cart.

Point of application moves, so force does work on mattress.

by the black dots in Figure 9.1). When you push against the brick wall, it neither moves nor deforms, and so the point of application of the force does not move. When you push the cart, you must move along with the cart to accelerate it; in other words, the point of application of the force is displaced. Similarly, when your force changes the shape of the deformable mattress in Figure 9.1*c*, the point of application of the force is displaced.

In order for a force to do work, the point of application of the force must undergo a displacement.

Whenever an external force is exerted on a system and the displacement of the point of application of the force is zero, no work is done by the force on the system. Conversely, in one dimension, when the **force displacement** (shorthand for "the displacement of the point of application of the force") is nonzero, the external force does work on the system, and so the energy of the system changes. (In more than one dimension, the angle between the force and the force displacement also plays a role. We shall discuss this point in Chapter 10.)* To distinguish the force displacement from other displacements, we add a subscript F to the symbol for displacement: $\Delta\vec{r}_F$.

EXERCISE 9.1 Displaced forces

For which of the following forces is the force displacement nonzero: (*a*) the force exerted by a hand compressing a spring, (*b*) the force exerted by Earth on a ball thrown upward, (*c*) the force exerted by the ground on you at the instant you jump upward, (*d*) the force exerted by the floor of an elevator on you as the elevator moves downward at constant speed?

SOLUTION (*a*), (*b*), and (*d*).

(*a*) The point of application of the force is at the hand, which moves to compress the spring.

(*b*) The point of application of the force of gravity exerted by Earth on the ball is at the ball, which moves.

(*c*) The point of application is on the ground, which doesn't move (at a macroscopic scale).

(*d*) The point of application is on the floor of the elevator, which moves. ✔

9.2 You throw a ball straight up in the air. Which of the following forces do work on the ball while you throw it? Consider the interval from the instant the ball is at rest in your hand to the instant it leaves your hand at speed v. (*a*) The force of gravity exerted by Earth on the ball. (*b*) The contact force exerted by your hand on the ball.

*Two additional external forces are exerted on the cart in Figure 9.1*b* that have nonzero force displacements: the force of gravity and the contact force exerted by the ground. However, these two forces are equal in magnitude and opposite in direction and so they "cancel out": Their net effect is the same as that of no force. Therefore no work is done by these forces on the cart.

9.2 Positive and negative work

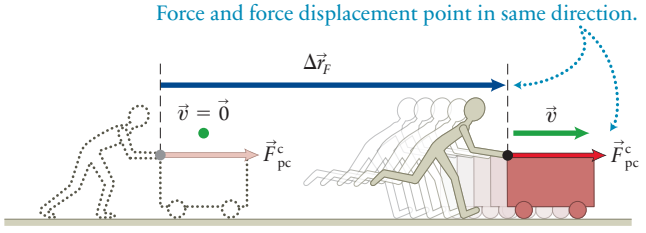
When the energy of a system increases as a result of an external force exerted on the system, the change in energy is positive, and so the work done by that external force on the system is said to be positive. An external force can also decrease the energy of a system, however. For instance, imagine slowing down a friend who is coasting on a bicycle. In this case, the kinetic energy of the system decreases, the change in energy is negative, and so the work done by the external force (that is, by you on your friend) on the system is said to be negative.

9.3 A ball is thrown vertically upward. (*a*) As it moves upward, it slows down under the influence of gravity. Considering the changes in energy of the ball, is the work done by Earth on the ball positive or negative? (*b*) After reaching its highest position, the ball moves downward, gaining speed. Is the work done by the gravitational force exerted on the ball during this motion positive or negative?

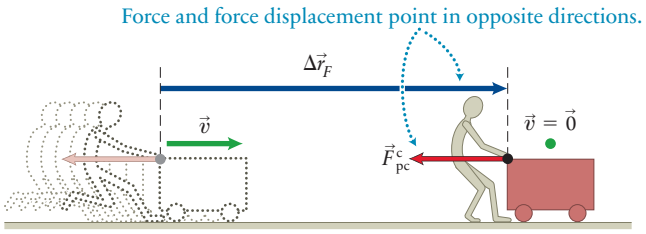
Examples of positive and negative work are illustrated in **Figure 9.2**. In both cases the person exerts a force $\vec{F}^{c}_{pc}$ on the cart. What, then, is the difference between these two situations? The difference lies in the direction in which the force is applied. To speed the cart up, the person must push in the direction of travel of the cart; the force vector $\vec{F}^{c}_{pc}$ and the force displacement $\Delta\vec{r}_F$ point in the same direction. To slow the cart down, the force must be applied in the direction opposite the cart's direction of travel; the force $\vec{F}^{c}_{pc}$ and the force displacement $\Delta\vec{r}_F$ point in opposite directions.

Figure 9.2 The work done on a system is positive if the system gains kinetic energy and negative if the system loses kinetic energy.

(*a*) Cart speeds up, so positive work is done on it

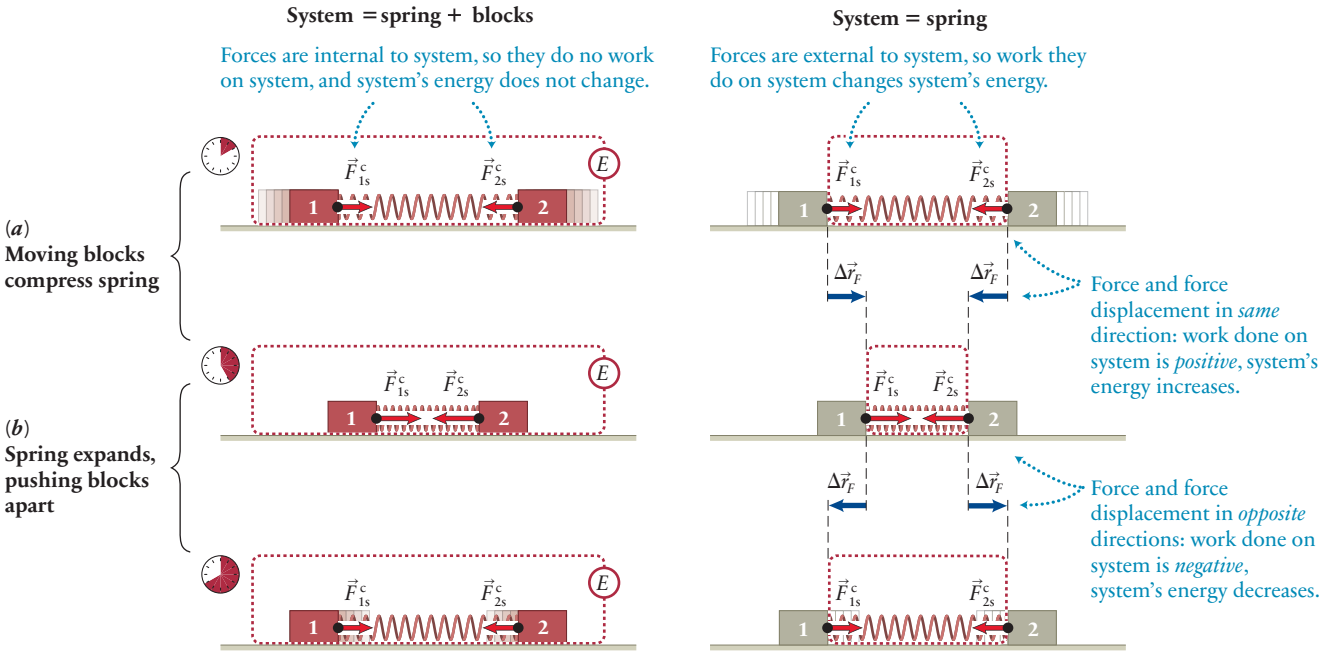


(*b*) Cart slows down, so negative work is done on it



CONCEPTS

Figure 9.3 Only forces that are *external* to a system can do work on the system and change the system's energy.



The work done by a force on a system is positive when the force and the force displacement point in the same direction and negative when they point in opposite directions.

9.4 Go back to Checkpoint 9.3 and answer the same two questions in terms of the directions of the applied forces and the force displacements. Do your answers agree with those you gave in Checkpoint 9.3?

Let us next consider a situation involving changes in potential energy. Figure 9.3*a* shows a spring being compressed by two moving blocks. There are two ways to look at this problem. We can consider the spring and the blocks as a closed system whose energy must remain constant, or we can consider just the spring as a system on which the blocks do work.

For the closed blocks-spring system in Figure 9.3*a*, the potential energy of the spring increases during the compression by an amount equal to the decrease in the blocks' kinetic energy. Because no external forces are exerted on the system, no work is involved.

Considering just the spring as our system in Figure 9.3*a*, we note that the increase in the spring's potential energy results from the work done by the blocks on it. Because the spring's potential energy increases, the work done on the system must be positive. Indeed, the forces exerted by the blocks on the spring point in the same direction as the force displacements.

Figure 9.3*b* shows the reverse situation: Two blocks are held against a compressed spring and then released, so that they accelerate away from each other. For the closed blocks-spring system, the energy doesn't change, and we can account for the increase in kinetic energy of the blocks by a decrease in the potential energy of the expanding spring. If, however, we consider the spring by itself as our system, the decrease in energy of the system implies that the work done by the blocks on the spring is negative.

Let us see if we reach the same conclusion for the sign of the work in Figure 9.3*b* when we look at the directions of the forces and the force displacements. Because we are interested in the work done *on the spring*, we need to consider the forces exerted by the blocks on it. These forces still point toward the spring, as shown in Figure 9.3*b*.* The force displacements, however, now point in the opposite direction, away from the spring. Because each force and its force displacement point in opposite directions, the work done by the blocks on the spring is negative. Everything is consistent: Both energy considerations and the directions of the forces and force displacements tell us that the work done by the blocks on the spring is negative as the spring expands.

CONCEPTS

*Careful! You may think the spring exerts a force on the blocks and not the other way around. Don't forget, however, that the force exerted by the spring on each block is just one part of an interaction pair. If you are having trouble picturing the force exerted by the blocks on the spring, imagine the spring expanding without having to push against the blocks: The expansion would be easier because there would be no resistance from the blocks. So, in a sense, the force exerted by the blocks on the spring slows down the expansion of the spring.

EXERCISE 9.2 Positive and negative work

Is the work done by the following forces considered in Exercise 9.1 positive, negative, or zero? In each case the system is the object on which the force is exerted. (*a*) The force exerted by a hand compressing a spring, (*b*) the force exerted by Earth on a ball thrown upward, (*c*) the force exerted by the ground on you at the instant you jump upward, (*d*) the force exerted by the floor of an elevator on you as the elevator moves downward at constant speed.

SOLUTION (*a*) Positive. To compress a spring, I must move my hand in the same direction as I push. ✔

(*b*) Negative. The force exerted by Earth points downward; the point of application moves upward. ✔

(*c*) Zero, because the point of application is on the ground, which doesn't move. ✔

(*d*) Negative. The force exerted by the elevator floor points upward; the point of application moves downward. ✔

9.5 Suppose that instead of the two moving blocks in Figure 9.3*a*, just one block is used to compress the spring while the other end of the spring is held against a wall. (*a*) Is the system comprising the block and the spring closed? (*b*) When the system is defined as being only the spring, is the work done by the block on the spring positive, negative, or zero? How can you tell? (*c*) Is the work done by the wall on the spring positive, negative, or zero?

9.3 Energy diagrams

In analyzing situations involving work, we extend the bar diagrams introduced in Chapter 5 to include work done on the system. Consider, for example, the situation in **Figure 9.4*a***. The energy of the system represented by this diagram increases from 10 J to 15 J. In order for this 5-J increase to occur, an amount of work equal to 5 J must be done by external forces on the system. The length of the vertical bar representing the final energy is equal to the sum of the lengths of the bars representing the initial energy and the work done by the external forces on the system.

Figure 9.4 Bar diagrams showing the work done on a system. Positive work adds to the system's energy; negative work subtracts from it.

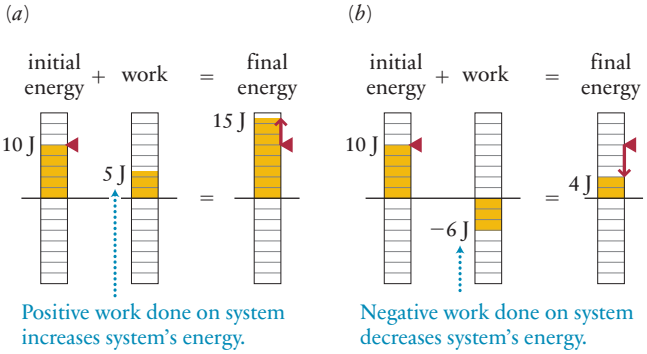


Figure 9.4*b* represents a system whose energy decreases. In order for the energy to decrease, the work done by external forces on the system must be negative, and so the bar representing this work extends below the baseline.

Because any of the four kinds of energy discussed in Chapter 7 (kinetic energy K, potential energy U, source energy E_s, and thermal energy E_{th}) can change in situations involving work, our bar diagrams need to include more detail in order to show in what form the energy transferred to a system appears when some external force does work on it. The additional bars make these diagrams more cumbersome to draw, however, and therefore we introduce the following simplification. Instead of drawing one set of bars for each of the initial and final individual energies, we draw just one set of bars representing the *change* in each category of energy, and we add a fifth bar, denoted by W, representing the work done by external forces on the system. We call these diagrams **energy diagrams.** Energy diagrams are a graphic representation of conservation of energy: They show that the energy of a system can change only due to the transfer of energy into or out of the system.

Just as with free-body diagrams, the first step in drawing an energy diagram is to define the system. In addition we should show any external force whose point of application undergoes a nonzero displacement; these are the forces that do work on the system. Consider, for example, the situation shown in **Figure 9.5*a***: a person pushing a pair of carts connected by a spring. The pushing accelerates the carts and compresses the spring. Let us choose a system comprising the pair of carts and the spring; Figure 9.5*a* shows its initial and final states. We also draw a vector representing the external force $\vec{F}^{c}_{pc}$ exerted by the person on the left cart and indicate the position of the point of application of that force with a black dot.

Because the carts speed up, the final kinetic energy of the system is greater than the initial kinetic energy and the change in kinetic energy $\Delta K = K_f - K_i$ is therefore positive. This change in kinetic energy corresponds to the upward-pointing arrow marked ΔK in Figure 9.5*b*, which is a set of the bar diagrams we used before now (no work values are shown). The compression of the spring causes the elastic potential energy stored in the spring to increase, and so the change in potential energy $\Delta U = U_f - U_i$ is also positive, as shown by the upward-pointing arrow in Figure 9.5*b*.

To represent this same information in an energy diagram (Figure 9.5*c*), we draw bars to represent the changes ΔK and ΔU. The heights of the bars, which are equal to the heights of the arrows in Figure 9.5*b*, now represent the *changes* in each energy category. The ΔE_s and ΔE_{th} bars are empty because there are no changes in these energies.

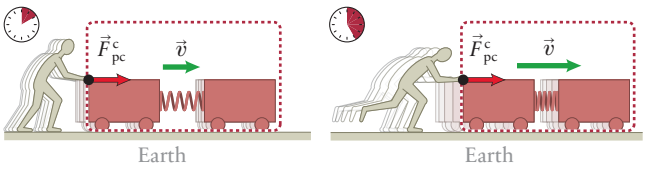
The change in the system's energy comes from the work done by external forces on the system. The length of the bar representing the work done on the system in Figure 9.5*c* must therefore equal the sum of the lengths of the ΔK and ΔU bars. Note that these energy arguments

CONCEPTS

Figure 9.5 (*a*) Changes in a system's energy can be shown (*b*) by initial and final bar diagrams or, more compactly, (*c*) by a single energy diagram, which shows the *changes* in energy in each category and a column showing work done on the system.

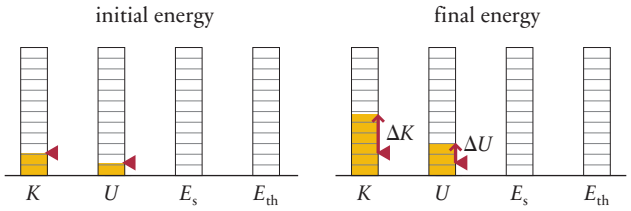
(*a*)

External work by person changes system's kinetic and potential energy.

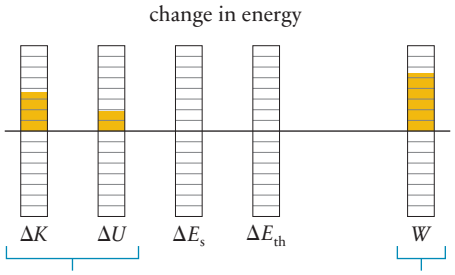


(*b*)

We can represent the changes in energy by initial and final bar diagrams . . .



(*c*) . . . or by a single **energy diagram**.



tell us that the work done by external forces on the system must be positive. We can verify this sign by checking the directions of the force exerted on the system and its force displacement. From Figure 9.5*a*, we see that both are to the right, and so the work done by external forces on the system is indeed positive.

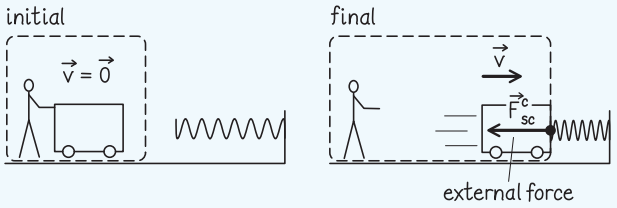
The general procedure for making energy diagrams is summarized in the Procedure box "Drawing Energy Diagrams"on page 288.

EXERCISE 9.3 Cart launch

A cart is at rest on a low-friction track. A person gives the cart a shove and, after moving down the track, the cart hits a spring, which slows down the cart as the spring compresses. Draw an energy diagram for the system that comprises the person and the cart over the time interval from the instant the cart is at rest until it has begun to slow down.

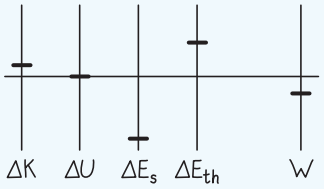
SOLUTION Following the steps in the Procedure box, I begin by listing the objects that make up the system: cart and person. Then I sketch the system in its initial and final states (**Figure 9.6**). The external forces exerted on the system are the force of gravity on person and cart, the contact force exerted by ground on person, the contact force exerted by track on cart, and the contact force exerted by spring on cart. The vertical forces add up to zero, and so the work done by them on the system is zero. The force exerted by the spring on the cart, however, undergoes a force displacement, so I include that force in my diagram and indicate its point of application, as shown in Figure 9.6.

Figure 9.6



Next I determine whether there are any changes in energy. Kinetic energy: The cart begins at rest and ends with nonzero speed, which means a gain in kinetic energy (**Figure 9.7**). Potential energy: Neither the configuration of the cart nor that of the person changes in a reversible way, so the potential energy does not change. (The configuration of the spring does change in a reversible way, but the spring is not part of the system.) Source energy: The person has to exert effort to give the cart a shove, so source energy is consumed. The source energy of the system therefore decreases. Thermal energy: No friction is involved, but the conversion of source energy is always accompanied by the generation of thermal energy, so the thermal energy increases. (The person's muscles heat up.)

Figure 9.7



In Figure 9.6 the force exerted by the spring on the cart points to the left and the force displacement points to the right. This means that the work done by the spring on the cart is negative, and the *W* bar of my energy diagram must extend below the baseline. I adjust the lengths of the bars so that the length of the *W* bar is equal to the sum of the lengths of the other bars, yielding the energy diagram shown in Figure 9.7. ✔

9.6 Draw an energy diagram for the cart in Figure 9.2*b*.

CONCEPTS

PROCEDURE: Drawing energy diagrams

1. Specify the **system** under consideration by listing the components inside the system.
2. **Sketch** the system in its initial and final states. (The initial and final states may be defined for you by the problem, or you may have to choose the most helpful states to examine.) Include in your sketch any external forces exerted on the system that undergo a nonzero force displacement, and draw a dot at the point of application of each force.
3. Determine any nonzero **changes in energy** for each of the four categories of energy, taking into account the four basic energy-conversion processes illustrated in Figure 7.13:
   - **a.** Did the speed of any components of the system change? If so, determine whether the system's kinetic energy increased or decreased and draw a bar representing ΔK for the system. For positive ΔK, the bar extends above the baseline; for negative ΔK, it extends below the baseline. (For some problems you may wish to draw separate ΔK bars for different objects in the system; if so, be sure to specify clearly the system component that corresponds to each bar and verify that the entire system is represented by the sum of the components.)
   - **b.** Did the configuration of the system change in a reversible way? If so, draw a bar representing the change in potential energy ΔU for the system. If necessary, draw separate bars for the changes in different types of potential energy, such as changes in elastic and gravitational potential energy.
   - **c.** Was any source energy consumed? If so, draw a bar showing ΔE_s. Source energy usually decreases, making ΔE_s negative, and so the bar extends below the baseline. Keep in mind that conversion of source energy is always accompanied by generation of thermal energy (Figure 7.13*c* and *d*).
   - **d.** Does friction occur within the system, or is any source energy consumed? If so, draw a bar showing ΔE_{th}. In nearly all cases we consider, the amount of thermal energy increases, and so ΔE_{th} is positive.
4. Determine whether or not any **work** W is done by external forces on the system. Determine whether this work is negative or positive. Draw a bar representing this work, making the length of the bar equal to the sum of the lengths of the other bars in the diagram.

   If no work is done by external forces on the system, leave the bar for work blank, then go back and adjust the lengths of the other bars so that their sum is zero.

CONCEPTS

EXERCISE 9.4 Compressing a spring

A block initially at rest is released on an inclined surface. The block slides down, compressing a spring at the bottom of the incline; there is friction between the surface and the block. Consider the time interval from a little after the release, when the block is moving at some initial speed v, until it comes to rest against the spring. Draw an energy diagram for the system that comprises the block, surface, and Earth.

SOLUTION I begin by listing the objects that make up the system: block, surface, and Earth. Then I sketch the initial and final states of the system (**Figure 9.8**). The spring exerts external forces on the system. The bottom end of the spring exerts a force on the surface edge, but this force has a force displacement of zero. The top end of the spring exerts a force $\vec{F}^{c}_{sb}$ on the block. Because this force undergoes a nonzero force displacement, I include it in my diagram and show a dot at its point of application.

Next I determine whether there are any energy changes. Kinetic energy: The block's kinetic energy goes to zero, and the kinetic energies of the surface and Earth do not change. Thus the kinetic energy of the system decreases, and ΔK is negative. Potential energy: As the block moves downward, the gravitational potential energy of the block-Earth system decreases, and so ΔU is negative. (Because the spring gets compressed, its elastic potential energy changes, but the spring is not part of the system.) Source energy: none (no fuel, food, or other source of energy is converted in this problem). Thermal energy: As the block slides, energy is dissipated by the friction between the surface and the block, so ΔE_{th} is positive.

To determine the work done on the system, I look at the external forces exerted on it. The point of application of the external force $\vec{F}^{c}_{sb}$ exerted by the spring on the block undergoes a force displacement opposite the direction of the force, so that force does negative work on the system. Thus the work done on the system by the external forces is negative, and the W bar extends below the baseline (**Figure 9.9**). I adjust the lengths of the bars so that the length of the W bar is equal to the sum of the lengths of the other three bars, yielding the energy diagram shown in Figure 9.9. ✔

Figure 9.8

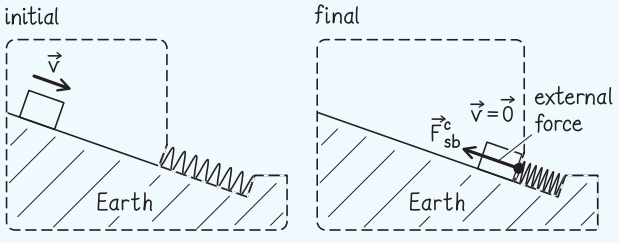
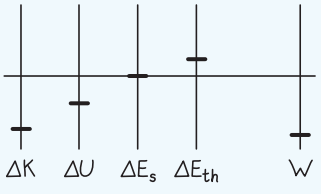


Figure 9.9



9.7 Draw an energy diagram for the situation presented in Exercise 9.4, but choose the system that comprises block, spring, surface, and Earth.

9.4 Choice of system

For a given situation, different choices of system yield different energy diagrams, as you just saw if you did Checkpoint 9.7. To examine this point further, consider the situation shown in **Figure 9.10**. Using a long rope, a person lowers a basket from a balcony, smoothly putting the basket down on the ground. The basket initially moves downward at speed v_i, but friction between the person's hands and the rope slows the basket so that it ends up at rest on the ground. For simplicity we assume no source energy is consumed; the person just lets the rope slide, exerting no physical effort.

Let us begin by choosing the basket and Earth as our system, a choice that leads to the energy diagram shown in Figure 9.10*a*. Because the basket comes to rest, its kinetic energy decreases. As the distance between Earth and basket decreases, the potential energy of the Earth-basket system decreases. There is no conversion of source energy nor any dissipation of energy within the system, and so we leave the columns for ΔE_s and ΔE_{th} blank.

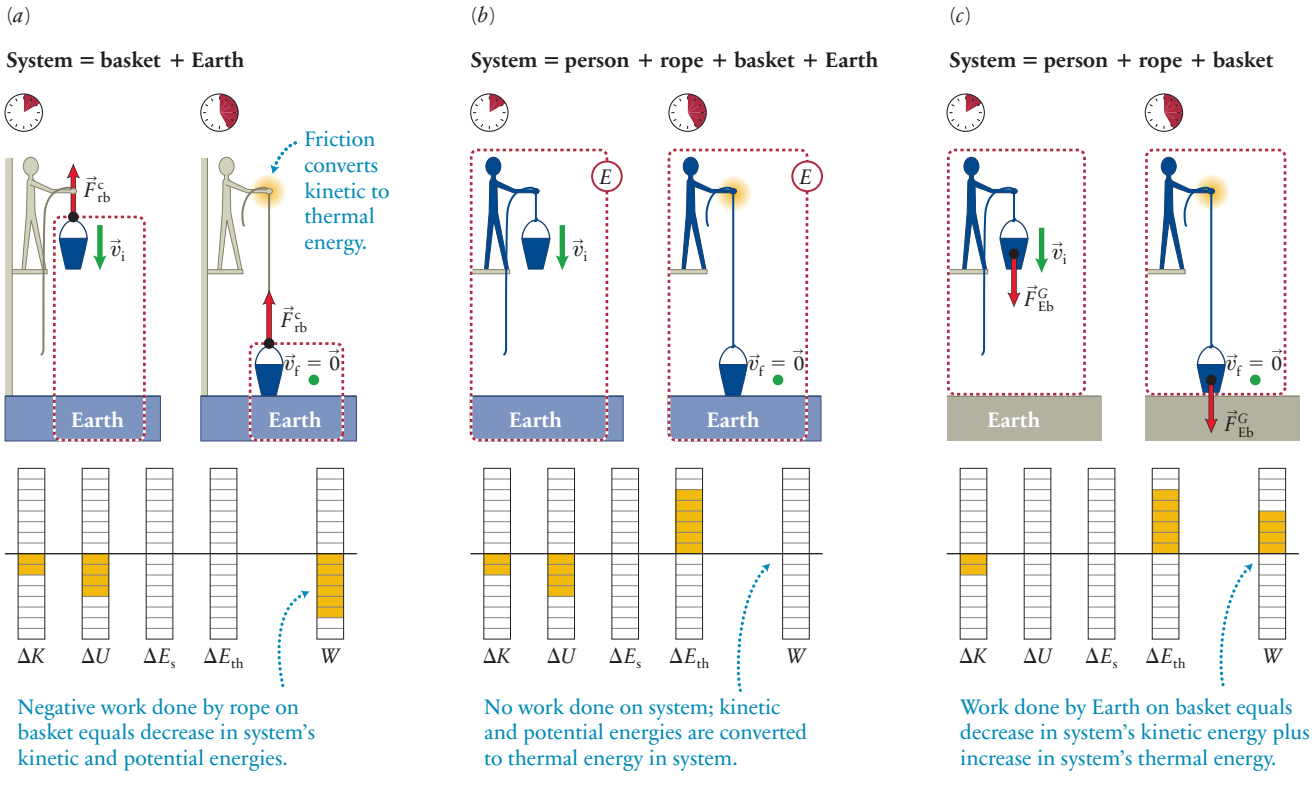
Next we examine the external forces exerted on the system. The person exerts forces both on Earth and, via the rope, on the basket. The latter involves a nonzero force displacement (the knot, which is the point of application for the force exerted by the rope on the basket, moves downward), so work is done on the system. This work is negative because the force exerted by the rope on the basket is upward and the force displacement is downward. In order to make the work equal to the change in the system's energy, the length of the bar representing work must be made equal to the sum of the lengths of the ΔK and ΔU bars.

What this energy diagram tells us is that the energy of the basket-Earth system decreases as energy is transferred out of the system. Because energy is transferred out of the system, the work done by the person on the system is negative.

What happens if we include the person and the rope in the system, as is Figure 9.10*b*? The changes in the kinetic and potential energies are the same as before, but now the person no longer exerts external forces and therefore cannot do any work on the system. In fact, because we have included everything in the system, there are no external forces and so there is no work done by external forces on the system—the system is closed.

For this choice of system, we have to take into account the dissipation of energy because of friction between the person's hands and the rope, and so we have a positive ΔE_{th}. The resulting energy diagram tells us that the initial kinetic energy of the basket and

Figure 9.10 Different choices of system yield different energy diagrams.



CONCEPTS

the initial potential energy of the system are converted to thermal energy (the rope and the person's hands heat up). No work is done by external forces on the system because the system is closed. This conclusion is entirely consistent with the one we drew on the basis of Figure 9.10*a*, but in the process we've obtained more information on where the energy went.

Finally we consider the same situation for a system that comprises only basket, person, and rope, as in Figure 9.10*c*. The changes in the kinetic and thermal energies are as before. For this choice of system, however, there is no change in gravitational potential energy (remember that gravitational potential energy refers to the configuration of Earth and the basket *together*, and Earth is no longer part of the system). What was counted as gravitational potential energy in Figure 9.10*a* and *b* appears as work done by the gravitational force on the system here. Earth exerts a downward external force on the basket as the basket moves downward, so Earth does positive work on the basket. (Earth also exerts a downward force on the person, but this force involves no force displacement.)

In Section 9.5, I shall show that the positive work done by the gravitational force on the system in Figure 9.10*c* is exactly equal to the negative of the change in the gravitational potential energy in Figures 9.10*a* and *b*. As this example shows, in order to avoid double-counting, it is important always to remember the following point:

Gravitational potential energy always refers to the relative position of various parts within a system, never to the relative positions of one component of the system and its environment.

Whenever you are looking at the energy associated with the gravitational interaction between an object inside your system and an object outside the system, do not include that energy in the ΔU bar of your energy diagram. That energy must be accounted for in the W part of the diagram instead.

The energy diagram of Figure 9.10*c* shows how the initial kinetic energy of the basket and the positive work done by Earth on the system end up as thermal energy.

All three choices of system lead to a valid conclusion. By looking at a situation from various points of view, you gain a better understanding of what energy conversions and transfers take place. Also, by looking at a situation in various ways and verifying that your answers are consistent with one another, you can convince yourself that you have not overlooked something.



9.8 Draw an energy diagram for just the basket in Figure 9.10.

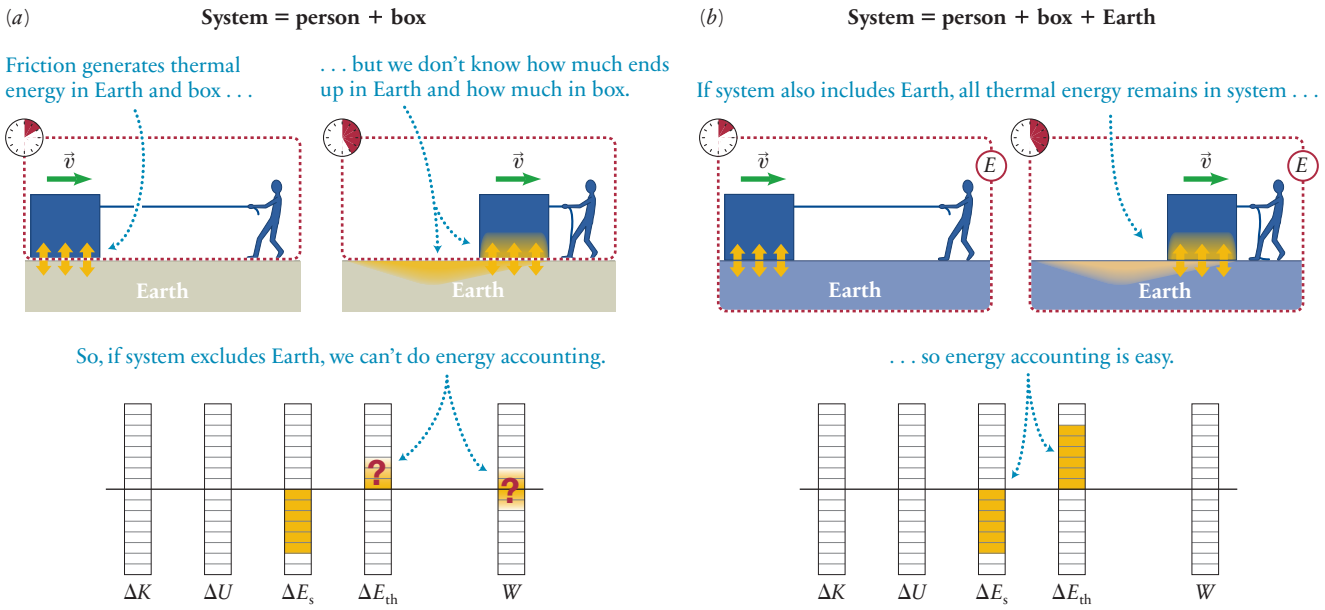
Although you are generally free to choose whichever system suits you best, there are certain choices of system that lead to complications. In particular:

When drawing an energy diagram, do not choose a system for which friction occurs at the boundary of the system.

The reason for this is simple: Wherever friction occurs, thermal energy is generated. The problem is that this thermal energy appears on both sides of the surface where friction occurs. If you pull on a rope to draw a large box toward yourself, as in **Figure 9.11**, friction between floor and box heats up both box and floor.

CONCEPTS

Figure 9.11 If you need to do energy accounting, don't choose a system for which friction occurs at the system boundary.



The thermal energy generated ends up partly in the floor and partly in the box. How much ends up where depends, among other things, on the materials that rub against each other and is not something that can be determined easily. So, if you choose just the box as your system, an unknown amount of energy flows out of the system, and it is no longer possible to do the energy accounting for the system.

There is another reason friction at a system boundary leads to problems. Whenever friction occurs at a boundary, the frictional force is *external* to the system (it is exerted by the environment on the system) and hence has to be taken into account in the computation of the work done on the system. The force of friction, however, is not applied at a single, well-defined location. Instead, it is distributed over the surfaces that move relative to each other. Consequently it is not possible to determine the force displacement, which is the other piece of information necessary to determine the work done on the system.

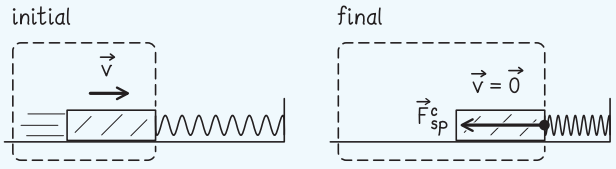
These problems with friction at a system boundary don't mean we can never treat any situations involving friction. Both our analyses of Figure 9.10 and Exercise 9.4 involve friction, but for all choices of system we considered, the two objects that rub against each other were kept together either inside or outside the system.

EXAMPLE 9.5 Puck and spring

A puck sliding on a horizontal surface hits the free end of a spring that is held fixed at the other end. There is friction between puck and surface. Consider the interval from the instant before the puck comes in contact with the spring until the instant the puck has zero velocity and the spring reaches maximum compression. Draw an energy diagram for a system that contains the puck and for which the work done on the system is nonzero. (Your system may include objects in addition to the puck if you like.)

1 GETTING STARTED I begin by making situation sketches at the beginning and end of the time interval of interest. The items in my sketches are the puck, the surface, and the spring (**Figure 9.12**).

Figure 9.12

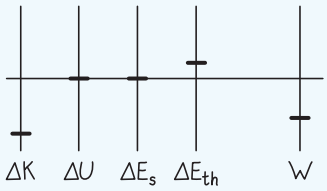


If I include all three items in my system, the system is closed, so the work done by external forces on the system is zero. Because there is friction between surface and puck, I must keep the two together. Given that the puck must be part of the system, I have no choice but to have my system be the puck plus the surface, so I draw my system boundary around these two objects. The only external force exerted on this system is the contact force exerted by the spring on the puck, so I add that force to my diagram and indicate its point of application.

2 DEVISE PLAN All I have to do is follow the procedure for drawing an energy diagram for my system.

3 EXECUTE PLAN My energy diagram is shown in **Figure 9.13**. Because the puck slows to a stop, its kinetic energy decreases. Because there are no reversible changes in the system configuration and because no source energy is consumed, my bars for ΔU and ΔE_s must be zero. The friction causes dissipation of energy, so the ΔE_{th} bar shows an increase.

Figure 9.13



Because the external force exerted by the spring on the system points in the direction opposite the direction of the force displacement, the work done by the spring on the system is negative. I adjust the lengths of the bars so that the length of the W bar is equal to the sum of the lengths of the other bars. ✔

4 EVALUATE RESULT The puck slows down under the combined action of the frictional force and the force exerted by the spring. So part of the puck's initial kinetic energy is converted to thermal energy, and the rest goes into compressing the spring. Because the spring is not part of the system, that energy leaves the system as (negative) work.

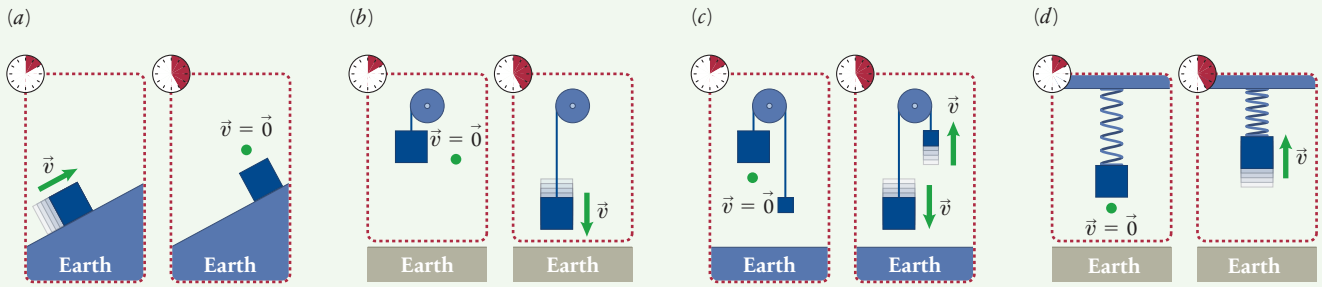
9.9 (*a*) Draw an energy diagram for the situation shown in Figure 9.11 for the system that comprises Earth and the box. Assume the box keeps moving at constant velocity, and consider the rope to be part of the person. (*b*) Explain each bar in the energy diagram of Figure 9.11*b*.

CONCEPTS

Self-quiz

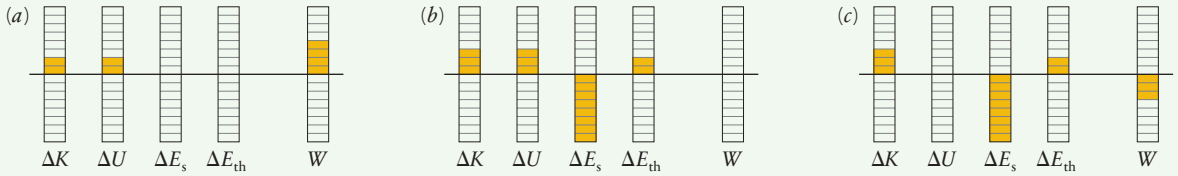
1. A stunt performer falls from the roof of a two-story building onto a mattress on the ground. The mattress compresses, bringing the performer to rest without his getting hurt. Is the work done by the mattress on the performer positive, negative, or zero? Is the work done by the performer on the mattress positive, negative, or zero?
2. Consider a weightlifter holding a barbell motionless above his head. (*a*) Is the sum of the forces exerted on the barbell zero? (*b*) Is the weightlifter exerting a force on the barbell? (*c*) If the weightlifter exerts a force, does this force do any work on the barbell? (*d*) Does the energy of the barbell change? (*e*) Are your answers to parts *c* and *d* consistent in light of the relationship between work and energy?
3. Do any of the systems in **Figure 9.14** undergo a change in potential energy? If yes, is the change positive or negative? Ignore any friction.

Figure 9.14



4. A person pushes a cart up a small hill, gaining speed along the way. Friction in the cart's wheels is negligible. Name the components of the system depicted in each energy diagram shown in **Figure 9.15**.

Figure 9.15



Answers

1. Because the mattress brings the performer to rest, $\vec{F}^{c}_{mp}$ must point upward. The point of application of this force moves downward as the mattress compresses, meaning the work done by the mattress on the performer is negative. The work done by the performer on the mattress is positive because $\vec{F}^{c}_{pm}$ points downward and its point of application moves downward.
2. (*a*) Yes, because the barbell remains motionless. (*b*) Yes. He exerts an upward force to counter the downward gravitational force. (*c*) No, because the point at which the lifter exerts a force on the barbell is not displaced. (*d*) No. (*e*) They are consistent. If no work is done on a system, the energy of the system does not change.
3. Yes. (*a*), (*c*), and (*d*). Because none of the situations involves friction or the conversion of source energy, the only things you need to consider are changes in the kinetic and potential energies and work done by external forces on the system. The system in part *a* is closed, and so the decrease in kinetic energy causes an increase in gravitational potential energy: $\Delta U > 0$. In part *b*, Earth is not included in the system, so there can be no potential energy component of the system's energy. In part *c*, the system is closed, so the increase in kinetic energy must be due to a decrease in gravitational potential energy: $\Delta U < 0$. In part *d*, Earth does negative work on the system, so the energy of the system must decrease. Because the kinetic energy increases, the elastic potential energy of the spring must decrease: $\Delta U < 0$.
4. If there is a change in potential energy, the cart and Earth must be in the system; if there is a change in source energy, the person must be in the system. (*a*) Cart and Earth. The person does positive work on the system. (*b*) Person, cart, and Earth. This system is closed; no work is done by external forces on it. Source energy is converted to increase the kinetic and potential energies (and some of it is dissipated). There is more kinetic and potential energy in this system than in part *a* because the person is included in the system. (*c*) Person and cart. Earth does negative work on the system.

9.5 Work done on a single particle

In the first part of this chapter we defined work as the change in the energy of a system due to external forces, and we represented work by the letter W. The SI unit of work is the same as that of energy, J. When work is done by external forces on a system, the energy of the system changes by an amount

$$\Delta E = W. \tag{9.1}$$

This equation is called the **energy law** and embodies conservation of energy: Because energy cannot be created or destroyed, the energy of a system can change only if energy is transferred into or out of the system.* For a system that is closed, $W = 0$ and so the energy law takes on the form we encountered in Chapter 5: $\Delta E = 0$.

To calculate the work done by an external force on a system, we start by considering a simple case: a single particle of inertia m subject to a constant force $\vec{F}$. The term **particle** refers to any object that has no internal structure and no extent in space. Because it lacks extent and internal structure, a particle cannot change shape and therefore any internal energy is fixed ($\Delta E_{\text{int}} = 0$). Only the kinetic energy of a particle can change, so

$$\Delta E = \Delta K \quad \text{(particle)}. \tag{9.2}$$

To calculate what happens to the kinetic energy of a particle subject to a constant force $\vec{F}$, we consider the particle's motion over a time interval $\Delta t = t_{\text{f}} - t_{\text{i}}$ for the case where the particle's velocity at instant t_{i} is $\vec{v}_{\text{i}}$. The constant force gives the particle a constant acceleration given by the equation of motion (Eq. 8.8):

$$a_x = \frac{\Sigma F_x}{m} = \frac{F_x}{m}. \tag{9.3}$$

At the end of the time interval, the x component of the particle's velocity is given by Eq. 3.4,

$$v_{x,\text{f}} = v_{x,\text{i}} + a_x \Delta t, \tag{9.4}$$

and the x component of the particle's displacement is given by Eq. 3.7,

$$\Delta x = v_{x,\text{i}} \Delta t + \tfrac{1}{2} a_x (\Delta t)^2. \tag{9.5}$$

Because we are considering a particle, the displacement of the particle is also the force displacement.

Using these expressions, we can evaluate the change in the particle's kinetic energy:

$$\Delta K = K_{\text{f}} - K_{\text{i}} = \tfrac{1}{2} m (v_{\text{f}}^2 - v_{\text{i}}^2). \tag{9.6}$$

*For now, we consider only mechanical transfers of energy and ignore any other type of energy transfer. Later we'll also take into account transfer of energy by heating and radiation.

Substituting Eq. 9.4 into Eq. 9.6 and keeping in mind that $v^2 = v_x^2$ and $a^2 = a_x^2$, we get

$$\begin{aligned}\Delta K &= \tfrac{1}{2}m[(v_{x,\mathrm{i}} + a_x\Delta t)^2 - v_{x,\mathrm{i}}^2] \\ &= \tfrac{1}{2}m[2v_{x,\mathrm{i}}a_x\Delta t + a_x^2(\Delta t)^2] \\ &= ma_x\left[v_{x,\mathrm{i}}\Delta t + \tfrac{1}{2}a_x(\Delta t)^2\right] \\ &= ma_x\Delta x_F = F_x\Delta x_F, \end{aligned} \tag{9.7}$$

where the substitution $v_{x,\mathrm{i}}\Delta t + \frac{1}{2}a_x(\Delta t)^2 = \Delta x_F$ comes from Eq. 9.5, the substitution $ma_x = F_x$ comes from Eq. 9.3, and Δx_F carries a subscript F to remind us that this is the **force displacement** (the displacement of the point of application of the force). Having found this result $\Delta K = F_x\Delta x_F$, we can use Eq. 9.2, $\Delta E = \Delta K$, to say that $\Delta E = F_x\Delta x_F$ and then Eq. 9.1, $\Delta E = W$, to finally say

$$W = F_x\Delta x_F \quad \text{(constant force exerted on particle, one dimension).} \tag{9.8}$$

In words:

For motion in one dimension, the work done by a constant force exerted on a particle equals the product of the x components of the force and the force displacement.

If the external force $\vec{F}$ and the force displacement $\Delta\vec{r}_F$ point in the same direction, F_x and Δx_F are either both positive or both negative, and so the product $F_x\,\Delta x_F$ is positive. Consequently, the kinetic energy increases. When the force and the particle's displacement point in opposite directions, F_x and Δx_F have opposite algebraic signs. The work done by the force on the particle is then negative, and so the kinetic energy of the particle decreases.

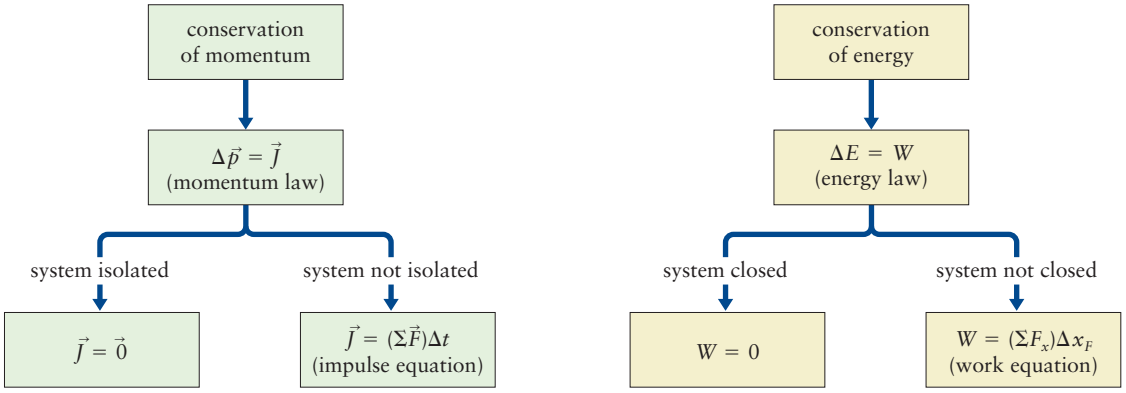
If more than one force is exerted on a particle, the x component of its acceleration is given by Eq. 8.8: $a_x = \sum F_x/m$. Substituting this expression for the x component of the acceleration into Eq. 9.7 gives $\Delta K = (\sum F_x)\Delta x_F$. Moving on to Eq. 9.8, we obtain the work done by these forces on the particle:

$$W = (\textstyle\sum F_x)\Delta x_F \quad \text{(constant forces exerted on particle, one dimension).} \tag{9.9}$$

Equation 9.9 is what we call the **work equation** for a particle. Together with the energy law (Eq. 9.1), this equation allows us to deal with systems that are not closed.

Notice the parallel between our treatments of momentum/impulse and energy/work (**Figure 9.16**). Conservation of momentum gives rise to the momentum law, Eq. 4.18, $\Delta\vec{p} = \vec{J}$, where $\vec{J}$ is the impulse delivered to the system. For an isolated system, the impulse is zero and so the momentum of the system does not change. If the system is not isolated, we can use the impulse equation, Eq. 8.25, $\vec{J} = (\sum\vec{F})\Delta t$, to calculate the change in the momentum of the system. Conservation of energy gives rise to the energy law, Eq. 9.1. The work done on a closed system is zero, and so the energy of the system does not change. If the system is not closed, we can use the work equation, Eq. 9.9, to calculate the change in the energy of the system. When we solve problems, it is possible to choose either a closed system or a system that is not closed. As the next example shows, both choices lead to the same answer. If you are unable to obtain a solution with one approach, try the other.

Figure 9.16 Energy and momentum changes for systems.



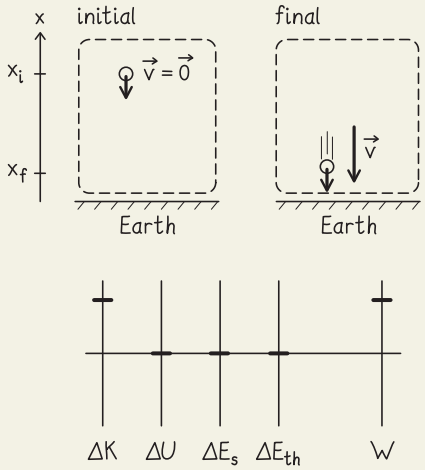
Although we derived Eqs. 9.8 and 9.9 for a particle, they can also be used for any rigid object as long as only the kinetic energy of the object changes. In that case, we can use Eq. 9.2 to calculate the change in the energy of the object. Because all particles that make up the object undergo the same displacement $\Delta\vec{r}$, Eq. 9.8 gives the work done by a constant force exerted on the object.

EXAMPLE 9.6 Work done by gravity

A ball of inertia m_b is released from rest and falls vertically. What is the ball's final kinetic energy after a displacement $\Delta x = x_f - x_i$?

1 GETTING STARTED I begin by making a sketch of the initial and final conditions and drawing an energy diagram for the ball (**Figure 9.17**). I choose an x axis pointing upward. Because the ball's internal energy doesn't change as it falls (its shape and temperature do not change), I can treat the ball as a particle. Therefore only its kinetic energy changes. I can also assume air resistance is small enough to be ignored, so that the only external force exerted on the ball is a constant gravitational force. This force has a nonzero force displacement and so does work on the ball. I therefore include this force in my diagram.

Figure 9.17



2 DEVISE PLAN If I treat the ball as a particle, Eqs. 9.1 and 9.2 tell me that the change in the ball's kinetic energy is equal to the work done on it by the constant force of gravity, the x component of which is given by Eq. 8.17: $F^G_{\text{Eb}\,x} = -m_b g$. (The minus sign means that the force points in the negative x direction.) To calculate the work done by this force on the ball, I use Eq. 9.8.

3 EXECUTE PLAN Substituting the x component of the gravitational force exerted on the ball and the force displacement $x_f - x_i$ into Eq. 9.8, I get

$$W = F^G_{\text{Eb}x}\Delta x_F = -m_b g(x_f - x_i).$$

Because the work is equal to the change in kinetic energy and the initial kinetic energy is zero, I have $W = \Delta K = K_f - 0 = K_f$, so

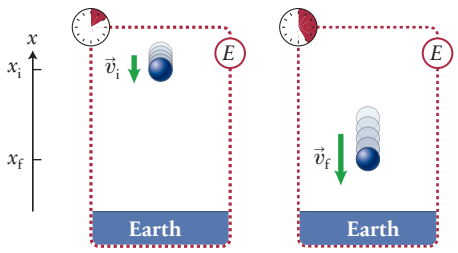
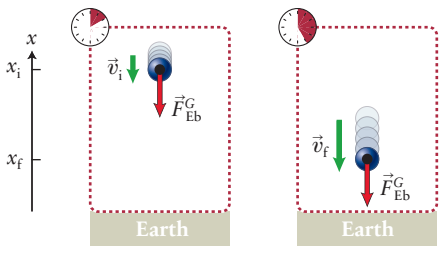
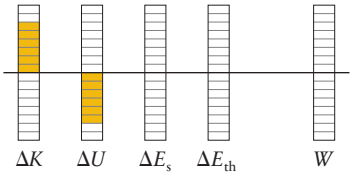
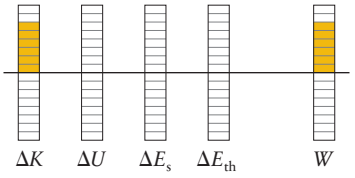
$$K_f = -m_b g(x_f - x_i). \checkmark$$

4 EVALUATE RESULT Because the ball moves in the negative x direction, $\Delta x = x_f - x_i$ is negative and so the final kinetic energy is positive (as it should be).

An alternative approach is to consider the closed Earth-ball system. For that system, the sum of the gravitational potential energy and kinetic energy does not change, and so, from Eq. 7.13, $\Delta K + \Delta U^G = \frac{1}{2}m_b(v_f^2 - v_i^2) + m_b g(x_f - x_i) = 0$. Because the ball starts at rest, $v_i = 0$, and so I obtain the same result for the final kinetic energy: $\frac{1}{2}m_b v_f^2 = -m_b g(x_f - x_i)$.

QUANTITATIVE TOOLS

Figure 9.18 Two ways to determine the change in kinetic energy of a falling ball.

	Closed-system approach	Open-system approach
Sketch		
Energy diagram		
Energy law	$\Delta E = 0$ $\Delta K + \Delta U^G = 0$	$\Delta E = \Delta K = W$ $W = F^G_{Ebx}\Delta x_F$
Equation to use	$\Delta U^G = m_b g(x_f - x_i)$	$F^G_{Ebx} = -m_b g$
Answer	$\Delta K = -m_b g(x_f - x_i)$	$\Delta K = -m_b g(x_f - x_i)$

The two approaches used in Example 9.6 are shown schematically in Figure 9.18. In the closed-system approach, the increase in kinetic energy is attributed to a decrease in the gravitational potential energy of the Earth-ball system. In the open-system approach, the increase in kinetic energy is attributed to the work done by the force of gravity on the ball. In this case it is meaningless to talk about gravitational potential energy because gravitational potential energy refers to the configuration of Earth and the ball *together*. Because Earth is not part of the system, it would not be clear whether to attribute this potential energy to Earth or to the ball.

9.10 Verify that in Example 9.6 the momentum transferred to the ball by the gravitational force is $\Delta p_x = m_b(v_{x,f} - v_{x,i})$.

9.6 Work done on a many-particle system

Everyday objects cannot always be treated as either single particles or rigid objects. As we saw in Section 8.12, we must consider such objects as systems of many interacting particles. The acceleration of the center of mass of such a system is given by Eq. 8.45:

$$\vec{a}_{cm} = \frac{\Sigma \vec{F}_{ext}}{m}, \tag{9.10}$$

where m is the inertia of the system. Using this result, we can now evaluate the work done by external forces on a many-particle system following an approach analogous to that of a single particle in Eqs. 9.4–9.8.

Consider the motion of a two-particle system over a time interval $\Delta t = t_f - t_i$. If the center of mass of the system moves at velocity $\vec{v}_{cm,i}$ at

QUANTITATIVE TOOLS

some initial instant t_i and has constant acceleration $\vec{a}_{cm}$, then from Eq. 9.4 the velocity of the center of mass at the end of the time interval is

$$\vec{v}_{cm,f} = \vec{v}_{cm,i} + \vec{a}_{cm}\Delta t \tag{9.11}$$

and its displacement is (from Eq. 9.5)

$$\Delta\vec{x}_{cm} = \vec{v}_{cm,i}\Delta t + \tfrac{1}{2}\vec{a}_{cm}(\Delta t)^2. \tag{9.12}$$

Using these expressions, we can evaluate the changes in the translational kinetic energy of the system from Eq. 6.32, $K_{cm} = \frac{1}{2}mv_{cm}^2$:

$$\Delta K_{cm} = K_{cm,f} - K_{cm,i} = \tfrac{1}{2}m(v_{cm,f}^2 - v_{cm,i}^2). \tag{9.13}$$

Following the same derivation used to obtain Eq. 9.7, we calculate for the change in translational kinetic energy of the system

$$\Delta K_{cm} = ma_{cmx}\Delta x_{cm} = (\Sigma F_{extx})\Delta x_{cm} \quad \text{(constant forces, one dimension)}. \tag{9.14}$$

Even though this result looks very much like Eq. 9.7, ΔK_{cm} does not represent the work done by the external force on the system. The reasons are that the translational kinetic energy K_{cm} is only part of the kinetic energy of the system (which we know from Eq. 6.35, $K = K_{cm} + K_{conv}$) and for a system of many interacting particles changes can also occur in the other kinds of energy. Therefore $\Delta K_{cm} \neq \Delta E$. Knowing from Eq. 9.1 that $\Delta E = W$, we see that ΔK_{cm} from Eq. 9.14 cannot be equal to the work done by the external force on the many-particle system:

$$\Delta K_{cm} \neq W \quad \text{(many-particle system)}. \tag{9.15}$$

To see this more explicitly, consider the system consisting of two carts connected by a spring, illustrated in **Figure 9.19**. Cart 1 is subject to a single external force $\vec{F}_{ext\,1}$. If we ignore any changes in the carts' internal energies, we can consider the carts to be particles, which means that all of the system's internal energy is stored in the spring. The pushing on cart 1 both accelerates the system and compresses the spring, causing changes in the system's kinetic and internal energies. Unfortunately, we cannot calculate these changes because we don't know how much the spring compresses and we don't know anything about the instantaneous velocities of the carts. Because the work done by external forces on the system is equal to the change in the system's energy E—where E is the sum of all forms of energy—Eq. 9.14 does not provide an expression for the work done on the system.

So, how can we determine the work done by external forces on a many-particle system? Because work is energy transferred, we can use the fact that the work W_{env} done by a system on its environment has the same magnitude as the work W_{sys} done by the environment on the system but carries the opposite sign: $W_{env} = -W_{sys}$. If the system gains, say, 100 J, then conservation of energy dictates that something outside the system must lose 100 J. In **Figure 9.20a**, for example, the external force $\vec{F}_{h1}$ exerted by the hand on cart 1 gives the cart a displacement $\Delta\vec{r}_F$. According to Eq. 8.15, $\vec{F}_{12} = -\vec{F}_{21}$, cart 1 must exert a force $-\vec{F}_{h1}$ on the hand. The work done by cart 1 on the hand over the displacement $\Delta\vec{r}_F$ is, from Eq. 9.8, $-F_{h1x}\Delta x_F$ (Figure 9.20*b*). Because the work done by the two-cart system *on* the hand is $-F_{h1x}\Delta x_F$, the work done *by* the hand on the two-cart

Figure 9.19 An external force does work on a system of two interacting carts.

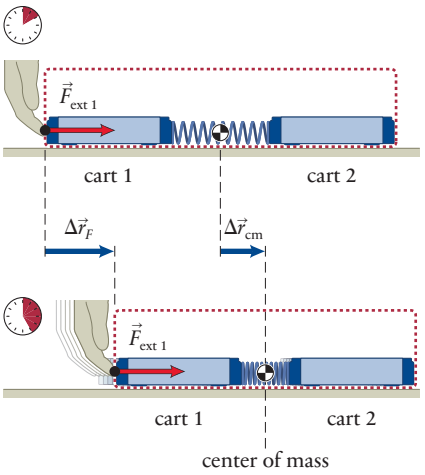
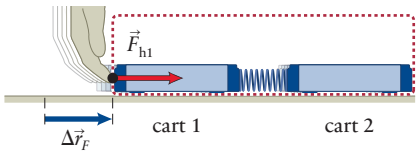
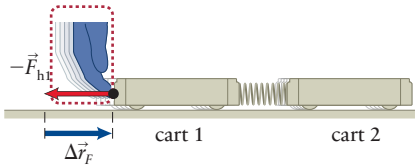


Figure 9.20 The work done by its environment on a system has the same magnitude as the work done by the system on its environment, but the two carry opposite signs: $W_{env} = -W_{sys}$.

(*a*) Hand does work on carts



(*b*) Carts do work on hand



system must be $+F_{\text{h1}\,x}\Delta x_F$. In general, when an external force is exerted on one of the particles in a two-particle system, the work equation is*

$$W = F_{\text{ext }1x}\Delta x_F \quad \text{(constant nondissipative force, one dimension).} \quad (9.16)$$

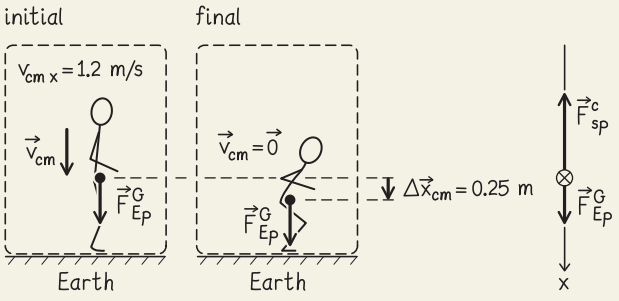
This work is equal to the product of the external force and Δx_F, the x component of the displacement *of the point of application of this force*, which is different from Δx_{cm}, the x component of the displacement *of the center of mass*, which appears in Eq. 9.14. Although we obtained this result for the specific case of two carts connected by a spring, the result does not depend on any of the specifics of the system and therefore is valid for any system of two interacting particles subject to a single, constant external force.

EXAMPLE 9.7 Landing on his feet

A 60-kg person jumps off a chair and lands on the floor at a speed of 1.2 m/s. Once his feet touch the floor surface, he slows down with constant acceleration by bending his knees. During the slowing down, his center of mass travels 0.25 m. Determine the magnitude of the force exerted by the floor surface on the person and the work done by this force on him.

❶ **GETTING STARTED** I begin by making a sketch of the initial and final conditions, choosing my person as the system and assuming the motion to be entirely vertical (**Figure 9.21**). I point the $+x$ axis downward in the direction of motion, which means that the x components of both the displacement and the velocity of the center of mass are positive: $\Delta x_{\text{cm}} = +0.25$ m and $v_{\text{cm}x} = +1.2$ m/s. Two external forces are exerted on the person: a downward force of gravity $\vec{F}^G_{\text{Ep}}$ exerted by Earth and an upward contact force $\vec{F}^c_{\text{sp}}$ exerted by the floor surface. Only the point of application of the force of gravity undergoes a displacement, and so I need to include only that force in my diagram. (Source energy and thermal energy balance each other; I ignore them here.)

Figure 9.21



❷ **DEVISE PLAN** Knowing the initial center-of-mass velocity, I can use Eq. 9.13 to calculate the change in the person's translational kinetic energy ΔK_{cm}. To calculate the vector sum of the external forces exerted on the person, I can use the value I obtain for ΔK_{cm} and the displacement $\Delta x_{\text{cm}} = +0.25$ m in Eq. 9.14 to obtain the vector sum of the forces exerted on the person. If I then subtract the force of gravity, I obtain the force exerted by the floor surface on the person. To determine the work done by this force on the person, I need to multiply it by the force displacement. Because the person slows down as he travels downward, his acceleration is upward and so the vector sum of the forces is upward too. To remind myself of this, I draw a free-body diagram in which the arrow for the upward contact force is longer than the arrow for the downward force of gravity.

❸ **EXECUTE PLAN** Because the person ends at rest, his final translational kinetic energy is zero, and Eq. 9.13 gives me

$$\Delta K_{\text{cm}} = 0 - \tfrac{1}{2}mv^2_{\text{cm,i}} = \tfrac{1}{2}(60\text{ kg})(1.2\text{ m/s})^2 = -43\text{ J}.$$

Substituting this value and the displacement of the center of mass into Eq. 9.14 yields

$$\Sigma F_{\text{ext}x} = \frac{\Delta K_{\text{cm}}}{\Delta x_{\text{cm}}} = \frac{-43\text{ J}}{0.25\text{ m}} = -170\text{ N}.$$

To obtain the force exerted by the floor from this vector sum, I add to the sketch a free-body diagram showing the two forces exerted on the person (Figure 9.21).

$$\Sigma F_{\text{ext}x} = F^G_{\text{Ep}x} + F^c_{\text{sp}x}$$

and so $F^c_{\text{sp}x} = \Sigma F_{\text{ext}x} - F^G_{\text{Ep}x}$. The x component of the force of gravity is $F^G_{\text{Ep}x} = mg = (60\text{ kg})(9.8\text{ m/s}^2) = +590$ N and so $F^c_{\text{sp}x} = -170\text{ N} - 590\text{ N} = -760$ N. ✔

To determine the work done by this force on the person, I must multiply the x component of the force by the force displacement. The point of application is at the floor, which doesn't move. This means that the force displacement is zero, and so the work done on the person is zero too: $W = 0$. ✔

❹ **EVALUATE RESULT** The contact force $F^c_{\text{sp}x}$ is negative because it is directed upward, as I expect. Its magnitude is larger than that of the force of gravity, as it should be in order to slow the person down.

*We can apply this line of reasoning only when the force $\vec{F}_{\text{ext }1}$ is nondissipative. As we saw in Section 9.4, if the force exerted at the boundary of a system is dissipative, we don't know how much of the dissipated energy ends up inside the system.

QUANTITATIVE TOOLS

9.11 Compare the rightmost terms of Eqs. 9.14 and 9.16 to determine whether, in any two-particle system subject to a single, constant external force, ΔE is larger than, equal to, or smaller than ΔK_{cm}. Explain your answer.

We can easily generalize Eq. 9.16 to a many-particle system subject to several constant forces. When more than one external agent interacts with the system, we can follow the same argument to calculate the work done by each external agent on the system. If agent 1 exerts a force $\vec{F}_{ext\ 1}$ on a system and the point of application of this force undergoes a displacement $\Delta\vec{r}_{F1}$, then the work done by that agent on the system is $W_1 = F_{ext\ 1x}\,\Delta x_{F1}$.* Likewise, the work done by agent 2 is $W_2 = F_{ext\ 2x}\,\Delta x_{F2}$. The work done by all the agents in the environment on the system is the sum of the work done by each on the system:

$$W = W_1 + W_2 + \cdots = F_{ext\ 1x}\Delta x_{F1} + F_{ext\ 2x}\Delta x_{F2} + \cdots \qquad (9.17)$$

or

$$W = \sum_n (F_{ext\ nx}\Delta x_{Fn} \quad \text{(constant nondissipative forces, one dimension)}, \qquad (9.18)$$

which is the work equation for a many-particle system subject to constant external forces.

Note that to calculate this work we must take the product of each external force and the force displacement *of that force*. If the system is an object that deforms, these displacement values are different in different parts of the system. For that reason both $F_{ext\,nx}$ and Δx_{Fn} in Eq. 9.18 carry a subscript n.

If the force is dissipative, we cannot use this expression because the dissipated energy is distributed on both sides of the system boundary. Dissipative forces are discussed in the next section. The results for a single particle and a system of particles are summarized in **Table 9.1**.

9.12 Show that for a one-particle system, Eqs. 9.14 and 9.18 both reduce to Eq. 9.9.

Table 9.1 Equation of motion and energy accounting for a particle and a system of particles

	Single particle or rigid object	System of particles or deformable object
Equation of motion	$\vec{a} = \dfrac{\Sigma\vec{F}}{m}$	$\vec{a}_{cm} = \dfrac{\Sigma\vec{F}_{ext}}{m}$
Energy law	$\Delta E = W$	$\Delta E = W$
Work equation	$W = (\Sigma F_x)\Delta x_F$	$W = \sum_n (F_{ext\,nx}\Delta x_{Fn})$
Kinetic energy change	$\Delta K = W$	$\Delta K_{cm} = (\Sigma F_{ext\,x})\Delta x_{cm}$

*When work is done by more than one force on an object or system, the work done by each force individually inherits the subscripts of the corresponding force.

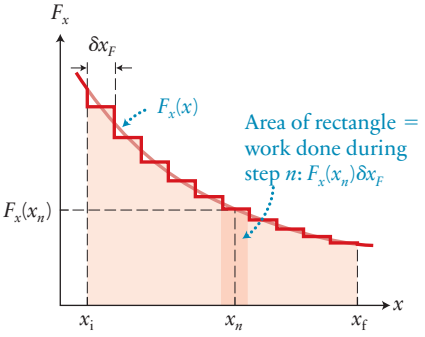
9.7 Variable and distributed forces

The results obtained in the preceding two sections apply only to forces that are constant and that are applied at well-defined locations in a system. If the forces are not constant or if we cannot determine the force displacement, then the results we obtained do not apply. Examples are the force exerted by a spring, which varies with the amount of compression or stretching, and the frictional force, which is distributed over a surface and for which it is not possible to determine the point of application. Let us examine these two cases in more detail.

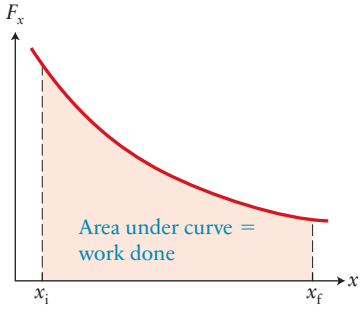
What should one do if the force exerted on a system isn't constant? (We call such a force a *variable force*.) For this kind of force, the x component can be written as a function of x: $F_x = F_x(x)$. The trick is to apply the constant-force result over a force displacement small enough to make the force approximately constant over that displacement. **Figure 9.22** shows how, for a force that varies with the position at which that force is applied, we can divide the force displacement into small segments.

Figure 9.22 Evaluating the work done by a variable force $\vec{F}(x)$ on a system.

(*a*) Work approximated by area under steps of constant force



(*b*) As δx_F approaches zero, area under curve approaches work done



Suppose we want to know the work done by the variable force on a system while the point of application of the force moves from x_i to x_f in Figure 9.22. We first divide the force displacement $x_f - x_i$ into N equal smaller force displacements δx_F such that the force is approximately constant across each δx_F (Figure 9.22*a*). Applying Eq. 9.8 to each smaller displacement δx_F, we can write the work done by the variable force on the system during each displacement as

$$W_n = F_x(x_n)\delta x_F, \tag{9.19}$$

where $F_x(x_n)$ is the x component of the force at x_n (this point x_n is located at the center of each interval, as shown in Figure 9.22*a*). This work is the area of the dark shaded rectangle in Figure 9.22*a*. The work done over the entire force displacement from x_i to x_f is then approximately equal to the sum of work W_n over all small force displacements δx_F between x_i and x_f:

$$W \approx \sum_n W_n = \sum_n F_x(x_n)\delta x_F, \tag{9.20}$$

where $\sum_n F_x(x_n)\delta x_F$ represents the sum of all the rectangular areas under the steps between x_i and x_f (the light shaded area in Figure 9.22*a*).

To get a better approximation, we can divide the displacement $x_f - x_i$ into a larger number of smaller force displacements δx_F, so that the variation of the external force over each displacement becomes even smaller and the red stepped curve is closer to the actual (light red) $F_x(x)$ curve.

An exact result is obtained if we let δx_F approach zero (and the number of small displacements approaches infinity), as in Figure 9.22*b*:

$$W = \lim_{\delta x_F \to 0} \sum_n F_x(x_n)\delta x_F, \tag{9.21}$$

which is precisely the definition of the integral of the x component of the external force with respect to x from x_i to x_f:

$$W = \int_{x_i}^{x_f} F_x(x)dx \quad \text{(nondissipative force, one dimension)}. \tag{9.22}$$

9.13 Show that Eq. 9.22 reduces to Eq. 9.8 when the force is constant.

An example of a force that varies with the position at which it is applied is the force exerted by a spring on a load: $\vec{F}_{sl} = -k(\vec{x} - \vec{x}_0)$ (Eq. 8.20).

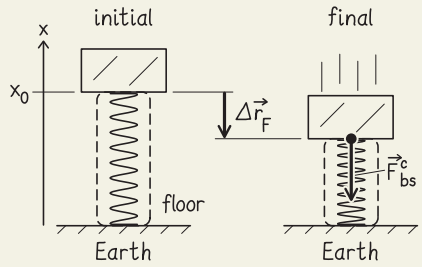
QUANTITATIVE TOOLS

EXAMPLE 9.8 Spring work

A brick of inertia m compresses a spring of spring constant k so that the free end of the spring is displaced from its relaxed position. What is the work done by the brick on the spring during the compression?

❶ **GETTING STARTED** I begin by making a sketch of the situation as the free end of the spring is compressed from its relaxed position x_0 to a position x (**Figure 9.23**). Because I need to calculate the work done by the brick on the spring, I choose the spring as my system. Three forces are exerted on the spring: contact forces exerted by the brick and by the floor, and the force of gravity. As usual when dealing with compressed springs, I ignore the force of gravity exerted on the spring (see Section 8.6). Only the force $\vec{F}^{c}_{bs}$ exerted by the brick on the spring undergoes a nonzero force displacement, so I need to show only that force in my diagram. Because the brick and spring do not exert any forces on each other when the spring is in the relaxed position, I do not draw this force in the initial state.

Figure 9.23



❷ **DEVISE PLAN** I need to calculate the work done by the brick on the spring. I could use Eq. 9.22 if I knew the force the brick exerts on the spring. That force is not given, but the force exerted by the brick on the spring and the force exerted by the spring on the brick form an interaction pair: $\vec{F}^{c}_{bs} = -\vec{F}^{c}_{sb}$. I can use Eq. 8.20 to determine $\vec{F}^{c}_{sb}$ and then use it in Eq. 9.22.

❸ **EXECUTE PLAN** Equation 8.20 tells me that the x component of the force exerted by the spring on the brick varies depending on how far the spring is compressed:

$$F_{sbx} = -k(x - x_0),$$

where x_0 is the coordinate of the relaxed position of the free end of the spring. The x component of the force exerted by the brick on the spring is thus

$$F_{bsx} = +k(x - x_0). \tag{1}$$

Because $x_0 > x$, F_{bsx} is negative, which means that $\vec{F}^{c}_{bs}$ points in the same direction as the force displacement. Thus the work done by the brick on the spring is positive. Now I substitute Eq. 1 into Eq. 9.22 and work out the integral to determine the work done by the brick on the spring:

$$W_{bs} = \int_{x_0}^{x} F_{bsx}(x)\,dx = \int_{x_0}^{x} k(x - x_0)\,dx$$

$$= \left[\tfrac{1}{2}kx^2 - kx_0x\right]_{x_0}^{x} = \tfrac{1}{2}k(x - x_0)^2. \checkmark \tag{2}$$

❹ **EVALUATE RESULT** Because the spring constant k is always positive (see Section 8.9 on Hooke's law), the work done by the brick on the spring is also positive. This is what I expect because the work done in compressing the spring is stored as potential energy in the spring.



9.14 In Example 9.8, consider the situation from the instant the brick is released to the instant at which it has zero velocity (when the spring compression is greatest). Draw an energy diagram for a system that comprises (*a*) the spring alone and (*b*) Earth, spring, and brick.

The energy law tells me that the change in the energy of the spring in Example 9.8 is equal to the work done by the brick on it: $\Delta E = W$. As you saw in part *a* of Checkpoint 9.14, only the potential energy of the spring (corresponding to the configuration of the spring) changes, and so $\Delta E = \Delta U_{spring} = W$. The change in the **potential energy of a spring** as its free end is displaced from its relaxed position x_0 to any position x is thus

$$\Delta U_{spring} = \tfrac{1}{2}k(x - x_0)^2. \tag{9.23}$$

This expression also describes the change in potential energy when the spring is stretched by some external force. To stretch the spring, the force exerted on it must point upward in Figure 9.23, which means that instead of $\Delta\vec{r}$ and $\vec{F}^{c}_{bs}$ both pointing downward as they do in compression, these two vectors both point upward. The expression for the work done by the stretching force on the spring (Eq. 2 in Example 9.8) remains the same because the displacement and the force exerted on the spring still both

point in the same direction and so Eq. 9.23 also remains the same. Consequently the potential energy of a spring increases both when the spring is compressed and when it is stretched.

9.15 (Brain twister) As the brick in Example 9.8 moves downward, why is the magnitude of the force it exerts on the spring given by Eq. 1 in Example 9.8 and not by simply the magnitude mg of the gravitational force exerted by Earth on the brick? Hint: Draw a free-body diagram for the brick as it moves downward.

Next we turn to the force of friction. As we discussed in Section 9.4, friction poses a problem for two reasons. First, the energy dissipated by friction ends up being divided between two objects. Second, there is no single point of application for frictional forces—the force is distributed over two surfaces that move relative to each other.

Consider a block sliding across a rough surface. The block is initially moving at some speed v but eventually comes to rest because of friction between it and the surface. **Figure 9.24** shows an energy diagram for the closed block-surface system. Because there are no reversible changes in configuration and no source energy is used, we know that all of the block's kinetic energy is converted to thermal energy. So, if we know the block's initial kinetic energy, we can determine the thermal energy generated from the energy law:

$$\Delta E = \Delta K + \Delta E_{\text{th}} = 0. \tag{9.24}$$

Because the final kinetic energy is zero, we have

$$\Delta E_{\text{th}} = -\Delta K = -(K_{\text{f}} - K_{\text{i}}) = K_{\text{i}}. \tag{9.25}$$

How are these energy changes related to the displacement $\Delta\vec{r}_{\text{cm}}$ over which the block slides to a stop? Let the force of friction exerted by the surface on the block be $\vec{F}^{\text{f}}_{\text{sb}}$. According to Eq. 9.14, the change in the block's translational kinetic energy is equal to the product of the external force exerted on the block and the displacement of its center of mass:

$$\Delta K_{\text{cm}} = F^{\text{f}}_{\text{sb}x}\Delta x_{\text{cm}}. \tag{9.26}$$

Because the frictional force and the block's displacement point in opposite directions, the right side of Eq. 9.26 is negative. Because the block is a rigid object, the translational kinetic energy is all of its kinetic energy, and so $\Delta K_{\text{cm}} = \Delta K$. Substituting Eq. 9.25 then yields an expression for the energy dissipated by friction:

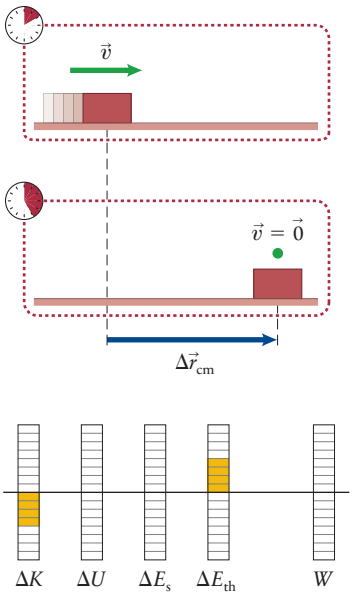
$$\Delta E_{\text{th}} = -F^{\text{f}}_{\text{sb}x}\Delta x_{\text{cm}} \quad \text{(constant dissipative force, one dimension)}. \tag{9.27}$$

When you use this equation, keep in mind that the energy is dissipated on both sides of the surface where the friction occurs. So, even though the right sides of Eqs. 9.26 and 9.27 are the product of a constant force and a displacement, it would not be correct to refer to this product as "the work done by the friction force exerted on the block." Furthermore, Eq. 9.27 holds for travel in only one direction; if the block reverses the direction of travel, you have to apply Eq. 9.26 to each one-way segment separately. If the sum of the distances traveled over all the one-way segments is d_{path}, the change in thermal energy becomes

$$\Delta E_{\text{th}} = F^{\text{f}}_{\text{sb}}d_{\text{path}} \quad \text{(constant dissipative force, one dimension)}, \tag{9.28}$$

where F^{f}_{sb} is the magnitude of the force of friction and where we have dropped the minus sign of Eq. 9.27 because all quantities are positive.

Figure 9.24 Energy diagram for a system in which friction dissipates kinetic energy to thermal energy.



QUANTITATIVE TOOLS

9.16 A 0.50-kg wood block slides 0.50 m on a horizontal floor before colliding elastically with a wall and reversing its direction of travel. If the block has an initial speed of 1.0 m/s and comes to rest exactly at its starting position, what is the magnitude of the frictional force between block and floor?

9.8 Power

Often we are interested not just in knowing how much energy is converted within a system or transferred to a system but also in knowing *how fast* the energy is converted or transferred. The rate at which energy is transferred or converted is called **power.** If the energy of a system changes by an amount ΔE over a time interval Δt, the average power is defined as

$$P_{\text{av}} = \frac{\Delta E}{\Delta t}. \tag{9.29}$$

The SI unit of power is the **watt** (W), which is equal to 1 J of energy per second: 1 W ≡ 1 J/s.* A person in good physical condition can deliver about 75 W of average power over a prolonged time interval. This means a healthy person can do work—shovel snow, accelerate an object, and so on—at the rate of about 75 J/s. An athlete can deliver about 400 W over a sustained time interval and up to about 1000 W during short time intervals.

The instantaneous power is

$$P = \lim_{\Delta t \to 0} \frac{\Delta E}{\Delta t} = \frac{dE}{dt}. \tag{9.30}$$

Equation 9.30 gives the rate at which the energy of a system changes. The concept of power can be applied to the rate at which any type of energy changes. For example, the power at which thermal energy is generated is given by

$$P = \frac{dE_{\text{th}}}{dt}. \tag{9.31}$$

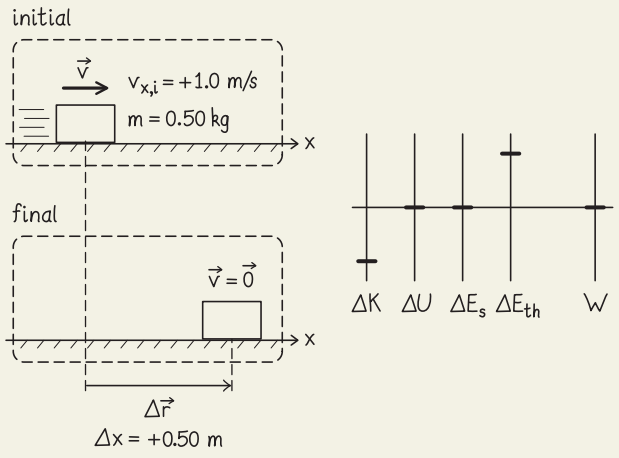
EXAMPLE 9.9 Frictional power

A 0.50-kg wood block initially traveling at 1.0 m/s slides 0.50 m on a horizontal floor before coming to rest. What is the average rate at which thermal energy is generated?

❶ **GETTING STARTED** I begin by making a sketch of the situation, listing all the given quantities and drawing an energy diagram for the closed block-floor system (**Figure 9.25**). I note that as the block slides to a stop, all of its kinetic energy is converted to thermal energy.

❷ **DEVISE PLAN** To calculate the average power at which thermal energy is generated, I need to know the amount ΔE_{th} of thermal energy generated and the time interval Δt over which it is generated. The ratio of these two quantities gives the average power (Eq. 9.29). The amount of thermal energy is equal to the initial kinetic energy of the block. If I assume the block has constant acceleration as it comes to a stop, I can use equations from Chapter 3 to determine Δt from the block's initial speed and the distance over which it comes to a stop.

Figure 9.25



(Continued)

*Be careful about distinguishing between italic *W*, our symbol for work, and nonitalic W, the abbreviation for watt, the SI unit of power.

QUANTITATIVE TOOLS

❸ **EXECUTE PLAN** Because the block ends at rest, I know from Eq. 3.4, $v_{x,f} = v_{x,i} + a_x \Delta t$, that

$$a_x = -\frac{v_{x,i}}{\Delta t}.$$

Substituting this result into Eq. 3.8, $x_f = x_i + v_{x,i}\Delta t + \frac{1}{2}a_x(\Delta t)^2$, gives

$$\Delta x = v_{x,i}\Delta t - \tfrac{1}{2}\frac{v_{x,i}}{\Delta t}(\Delta t)^2 = \tfrac{1}{2}v_{x,i}\Delta t$$

or

$$\Delta t = \frac{2\Delta x}{v_{x,i}} = \frac{2(0.50\text{ m})}{1.0\text{ m/s}} = 1.0\text{ s}.$$

The block's initial kinetic energy is $\frac{1}{2}mv_{x,i}^2 = \frac{1}{2}(0.50\text{ kg})(1.0\text{ m/s})^2 = 0.25\text{ J}$, so the thermal energy generated is $\Delta E_{th} = 0.25$ J. From Eq. 9.29, I then obtain for the average power at which this energy is generated: $P_{av} = (0.25\text{ J})/(1.0\text{ s}) = 0.25\text{ W}$. ✔

❹ **EVALUATE RESULT** The value I obtain is rather small—as a gauge, consider that a 25-W light bulb consumes energy at 100 times this rate—but it's unlikely that the energy of a sliding block would be sufficient to generate any light.

Next consider a situation where a system's energy is changed by a constant external force $\vec{F}_{ext}$. One such instance might be when you push a cart. As long as the force you exert on the cart is constant, we can analyze the power being supplied by you to the cart in terms of the x components of the force $F_{ext\,x}$ and the force displacement Δx. We know from Eqs. 9.1 and 9.8 that

$$\Delta E = W = F_{ext\,x}\Delta x. \tag{9.32}$$

Substituting this expression for ΔE into Eq. 9.29 yields

$$P_{av} = \frac{\Delta E}{\Delta t} = F_{ext\,x}\left(\frac{\Delta x}{\Delta t}\right) \tag{9.33}$$

or, with the definition of average velocity,

$$P_{av} = F_{ext\,x}v_{x,av}, \tag{9.34}$$

where $v_{x,av}$ is the x component of the average velocity of the point of application of the force (and because the cart is a rigid object it is also the x component of the average velocity of the cart). The power is positive when $\vec{F}_{ext}$ and $\vec{v}_{av}$ point in the same direction, and in this case the energy of the system increases. When $\vec{F}_{ext}$ and $\vec{v}_{av}$ point in opposite directions, the power is negative and the energy of the system decreases.

To obtain the instantaneous power, we take the limit as Δt in Eq. 9.33 goes to zero, yielding

$$P = F_{ext\,x}v_x \quad \text{(one dimension)}. \tag{9.35}$$

EXAMPLE 9.10 Car drag

A car requires 300 kJ of energy to overcome air resistance and maintain a constant speed of 20 m/s over a distance of 1.0 km. What is the force of air resistance exerted on the car?

❶ **GETTING STARTED** The air resistance causes dissipation of the car's source energy. The problem statement tells me that 300 kJ is dissipated over a distance of 1.0 km when the car moves at 20 m/s.

❷ **DEVISE PLAN** From the speed and the distance traveled, I can calculate the time interval over which the 300 kJ is required. I can then use Eq. 9.29 to calculate the rate at which the energy is converted. Knowing the velocity, I can then obtain the force of air resistance from Eq. 9.35.

❸ **EXECUTE PLAN** At 20 m/s it takes the car 50 s to cover 1.0 km. The rate at which energy is converted is thus $P = (300\text{ kJ})/(50\text{ s}) = 6.0\text{ kW}$. From Eq. 9.35, I obtain

$$F_x = \frac{P}{v_x} = \frac{6.0\text{ kW}}{20\text{ m/s}} = 300\text{ N}. \checkmark$$

QUANTITATIVE TOOLS

4 EVALUATE RESULT The answer I obtain—300 N—is the magnitude of the gravitational force exerted on a 30-kg (65-lb) object and therefore equal in magnitude to the force required to hold up such an object. I know, however, that if I stick my hand out a car window at highway speed, the force exerted by the air on my hand is not much larger than the force required to hold up a small object. The car, being much larger than my hand, intercepts a lot more air, but it is shaped more aerodynamically, and so it is reasonable that the force of air resistance is only 300 N.

9.17 (*a*) A gallon of gasoline contains about 1.4×10^8 J of chemical energy. A car consumes this amount of gasoline in approximately 30 min when cruising along a highway; a plane consumes the same amount in about 1 s when flying at cruising altitude. What is the average power of the energy release in each case? (*b*) A 10-kg load must be hoisted up the side of a building that is 50 m tall. How quickly can an athlete capable of delivering 500 W of power get the job done? (*c*) A 3-kW engine gets the job done six times faster than the athlete. How does the work done by the engine on the load compare with the work done by the athlete on the load?

Chapter Summary

Work done by a constant force (Sections 9.1, 9.2, 9.5, 9.6)

Concepts In order for a force to do work on an object, the point of application of the force must undergo a displacement.

The SI unit of work is the **joule** (J).

The work done by a force is positive when the force and the force displacement are in the same direction and negative when they are in opposite directions.

Quantitative Tools When one or more constant forces cause a particle or a rigid object to undergo a displacement Δx in one dimension, the work done by the force or forces on the particle or object is given by the **work equation:**

$$W = (\Sigma F_x)\Delta x_F. \tag{9.9}$$

In one dimension, the work done by a set of constant nondissipative forces on a system of particles or on a deformable object is

$$W = \sum_n (F_{\text{ext}\,nx}\Delta x_{Fn}). \tag{9.18}$$

If an external force does work W on a system, the **energy law** says that the energy of the system changes by an amount

$$\Delta E = W. \tag{9.1}$$

For a closed system, $W = 0$ and so $\Delta E = 0$.

For a particle or rigid object, $\Delta E_{\text{int}} = 0$ and so

$$\Delta E = \Delta K. \tag{9.2}$$

For a system of particles or a deformable object,

$$\Delta K_{\text{cm}} = (\Sigma F_{\text{ext}\,x})\Delta x_{\text{cm}}. \tag{9.14}$$

Energy diagrams (Sections 9.3, 9.4)

Concepts An **energy diagram** shows how the various types of energy in a system change because of work done on the system. (See the Procedure box on page 288.)

In choosing a system for an energy diagram, avoid systems for which friction occurs at the boundary because then you cannot tell how much of the thermal energy generated by friction goes into the system.

Variable and distributed forces (Section 9.7)

Concepts The force exerted by a spring is variable (its magnitude and/or direction changes) but nondissipative (no energy is converted to thermal energy).

The frictional force is dissipative and so causes a change in thermal energy. This force is also a distributed force because there is no single point of application.

Quantitative Tools The work done by a variable nondissipative force on a particle or object is

$$W = \int_{x_i}^{x_f} F_x(x)dx. \tag{9.22}$$

If the free end of a spring is displaced from its relaxed position x_0 to position x, the change in its potential energy is

$$\Delta U_{\text{spring}} = \tfrac{1}{2}k(x - x_0)^2. \tag{9.23}$$

If a block travels a distance d_{path} over a surface for which the magnitude of the force of friction is a constant F^{f}_{sb}, the energy dissipated by friction (the thermal energy) is

$$\Delta E_{\text{th}} = F^{\text{f}}_{\text{sb}} d_{\text{path}}. \tag{9.28}$$

Power (Section 9.8)

Concepts Power is the *rate* at which energy is either converted from one form to another or transferred from one object to another.

The SI unit of power is the **watt** (W), where 1 W = 1 J/s.

Quantitative Tools The **instantaneous power** is

$$P = \frac{dE}{dt}. \tag{9.30}$$

If a constant external force $F_{\text{ext}x}$ is exerted on an object and the x component of the velocity at the point where the force is applied is v_x, the power this force delivers to the object is

$$P = F_{\text{ext}x}v_x. \tag{9.35}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

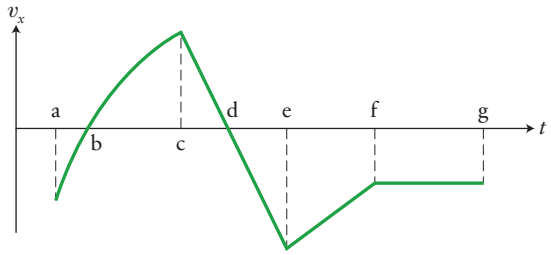
9.1 Force displacement

1. If a large force is applied to an object, does it necessarily follow that work in done on it? •
2. Hitting a door with your bare fist hurts more than hitting a sofa cushion. In work terms, explain why this is so. •
3. If you drop a brick from a height of 50 mm onto your toe, it probably won't hurt much, but if you drop the brick from a height of 0.5 m, it will hurt. The force of gravity exerted on the brick is the same in both cases. Explain why it hurts more from the increased height. ••
4. An object moves with constant velocity. What can you say about the work done on a system that includes only this object? ••
5. You hold a small steel ball in your hand above a laundry basket. Then you throw it straight up and watch as it rises, falls, and comes to rest in the pile of dirty laundry. Ignoring friction, discuss the forces exerted on the ball, the force displacements associated with those forces, and the work done by each force during this process. ••

9.2 Positive and negative work

6. The x component of velocity of a particle as a function of time is shown in Figure P9.6. Over what intervals is the work done on the particle (*a*) positive, (*b*) negative, (*c*) zero? •

Figure P9.6

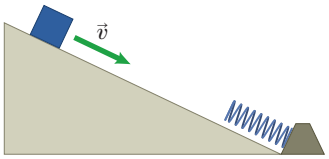


7. You try to push a heavy box along the floor while you are dressed in socks, but your feet keep slipping. When you put on some shoes, your feet no longer slip, and you are able to do work on the box as you push it. Has the friction force between your shoes and the floor done positive or negative work? •
8. You are in an elevator moving between two floors of a building at constant speed. Compare the work done on you by the elevator when you are moving upward to a higher floor to the work done when you are moving downward to a lower floor. ••
9. As a car accelerates from rest to highway speed, a pedestrian watching the motion sees positive work done on the car because its kinetic energy increases. However, from the reference frame of someone driving in the same direction at highway speed, the car is initially moving (backward) and then becomes stationary, corresponding to negative work. Which observer is right? ••
10. A system consists of two interacting particles 1 and 2, with particle 1 confined to an x axis and the system in equilibrium when particle 1 is at a given position on that axis. If displacing particle 1 in either direction from this position requires positive work to be done on the system by an external force, what is the shape of the potential energy curve (as a function of position) for the system in the region near equilibrium? •••

9.3 Energy diagrams

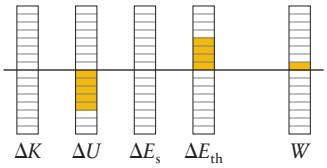
11. Give examples of cases where work is done on a system without changing its kinetic energy. •
12. Taking off on a runway, a small jet experiences a forward force of 90,000 N exerted by the engine and a backward force of 16,000 N exerted by the air. Choose a system and draw an energy diagram using the beginning and end of the runway as your initial and final positions. ••
13. A block launched at speed v down an incline slides down the incline (there is friction between the block and incline) and into a spring (Figure P9.13), where it slows to zero speed as the spring compresses. (*a*) Draw an energy diagram for this process using a system consisting of the block, spring, Earth, and incline. If there is more than one object causing energy conversions, use a separate bar for each. (*b*) Repeat part *a* using a system of just the block and incline. (*c*) Why can't you use the block alone as your system? ••

Figure P9.13



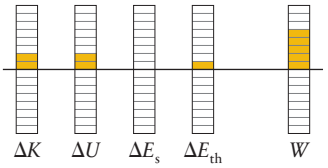
14. What, if anything, is wrong in the energy diagram in Figure P9.14 for a system comprising Earth, an inclined plane, and a block sliding along the incline? The block slides down the incline at constant speed. The only forces exerted on the block are the force of gravity and friction with the incline. ●●

Figure P9.14



15. What possible scenario is described by the energy diagram in Figure P9.15? ●●●

Figure P9.15



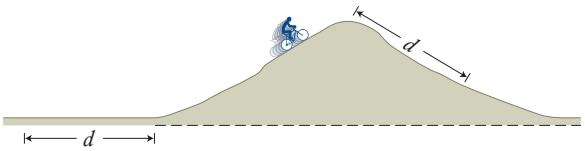
9.4 Choice of system

16. You are lifting a ball at constant velocity. (*a*) When the system is the ball, is work done on the system? If so, by what agent(s)? (*b*) Describe the potential energy of this system during the lift. (*c*) When the system comprises the ball and Earth, is work done on the system? If so, by what agent? (*d*) Describe how the potential energy of the ball-Earth system changes during the lift. ●
17. You throw a ball straight up into the air. If air resistance affects the motion of the ball, which takes longer: the upward trip or the downward trip? Analyze using three different systems. ●
18. A tomato is launched vertically from a spring-loaded toy cannon placed on the pavement. The tomato rises to some maximum height and then falls back, but you have removed the cannon so that the tomato splats on the pavement. Discuss the energy conversions during this process using the following systems: (*a*) tomato, cannon, air, Earth, pavement; (*b*) tomato, cannon, air, pavement; (*c*) tomato, air, pavement; and (*d*) tomato. ●●
19. A ball in motion collides elastically with an identical ball at rest. If we treat each ball as a separate system, is the work done on each ball the same? ●●

9.5 Work done on a single particle

20. You throw a 100-g ball upward with a speed of 19.8 m/s. How much work does the force of gravity do on the ball during its trip to its maximum height? ●
21. At the end of a delivery ramp, a skid pad exerts a constant force on a package so that the package comes to rest in a distance d. When the pad is replaced by one that requires a distance of $2d$ to stop the same package, what happens to the length of time it takes for the package to stop? ●
22. At the end of a delivery ramp, a skid pad exerts a constant force on a package so that the package comes to rest in a distance d. The ramp is changed so that the same package arrives at the skid pad at a higher speed and the stopping distance is $2d$. What happens to the time interval required for the package to stop? ●
23. A 48-kg acrobat must jump high and land on his brother's shoulders. To accomplish this, she leaps from a crouched position to a height where her center of mass is 1.60 m above the ground. Her center of mass is 320 mm above the ground in the crouch and 900 mm above the ground when her feet leave the ground. (*a*) What is the average force exerted on her by the ground during the jump? (*b*) What is her maximum speed? ●●
24. A beetle that has an inertia of 4.0×10^{-6} kg sits on the floor. It jumps by using its muscles to push against the floor and raise its center of mass. (*a*) If its center of mass rises 0.75 mm while it is pushing against the floor and then continues to travel up to a height of 300 mm above the floor, what is the magnitude of the force exerted by the floor on the beetle? (*b*) What is the beetle's acceleration during the motion? ●●
25. A bicyclist pedals a distance d at constant speed on a level stretch of road. She then pedals up a hill and passes slowly over the top. As she coasts down the other side of the hill, she applies the brakes and comes to a stop when she is a distance d past the top (Figure P9.25). Choose an appropriate system and draw an energy diagram from the instant the bicyclist begins her climb until the instant she has stopped on the downslope. ●●

Figure P9.25

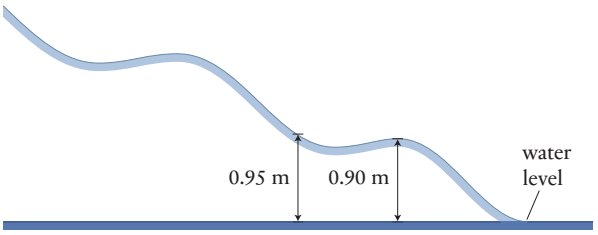


26. In emergency braking, a certain car requires 7.0 m to come to a stop from an initial speed of 10 m/s. Use an argument based on work done on the car to determine the braking distance required if the car's initial speed is 30 m/s. ●●
27. At instant t_0, you are sitting on a dock and you begin using a rope to lower a 4.0-kg lobster trap into the water, 1.4 m below. You lower the trap at a constant speed of 1.0 m/s. At instant t_1 it just reaches the water surface 1.4 m below the dock. You slacken the rope and the trap continues to move downward at a constant speed of 1.0 m/s until at instant t_2 the trap reaches the bottom of the 10-m-deep bay. Choose a system that includes the water, Earth, and the trap. (*a*) Do you do any work on the system in the interval from t_0 to t_2? (*b*) What forces are exerted on the trap during the interval from t_0 to t_2? (*c*) Draw an energy diagram for the interval from t_0 to t_2, showing a correct numerical scale. ●●

28. A 1.1-kg lobster climbs into the trap of Problem 27, and at instant t_3 you begin to haul the trap up onto the dock. The trap moves at a constant 0.40 m/s through the water, and it breaks the surface of the water at instant t_4. You adjust your pull so that it continues moving at that speed until it reaches the height of the dock at instant t_5. Use a system of water, Earth, trap, and lobster. (*a*) Do you do any work on the system in the interval from t_3 to t_5? (*b*) What forces are exerted on the trap as it rises? (*c*) Draw an energy diagram for this system from instant t_3 to instant t_5. In your diagram, there are two kinds of energy you have to combine because you have no way to separate them from each other. Indicate this in your labels. ●●

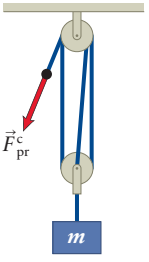
29. A 30-kg child slides down the amusement-park wave slide shown in Figure P9.29. When they are 0.95 m above the water level their speed is 2.2 m/s. At the top of the final hump, at a height of 0.90 m above the water level, their speed is also 2.2 m/s. What is the increase in the thermal energy of the child and the slide as they move between these two points? ●●

Figure P9.29



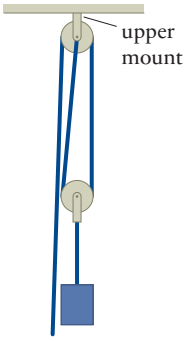
30. In the block-and-tackle arrangement shown in Figure P9.30, three segments of the single rope pull on the block. (*a*) Show that the magnitude F_{pr}^{c} of the force exerted by a person on the rope to raise the block at constant speed is $mg/3$, where m is the inertia of the block. (*b*) One worker uses this arrangement to lift a heavy block, while another hauls up an identical block with a straight rope. After both blocks have been raised to the same second-floor window, has either worker done more work than the other? If so, which worker has done more work? ●●

Figure P9.30



31. The pair of pulleys shown in Figure P9.31 is used to lift a 50-kg block 0.25 m. The constant acceleration of the block during the lift is +2.5 m/s^2, and the pulleys and rope have negligible inertia and negligible friction. What is the tension (*a*) in the pulling rope and (*b*) in the rope from which the block hangs? (*c*) What force is exerted by the upper mount on the top pulley? (*d*) How much work is done by the person pulling the rope in lifting the block 0.25 m? (*e*) By how much does the kinetic energy of the block change in this process? (*f*) How much work is done on the block by the rope attached to it? (*g*) How much work does the ceiling do in this process? (*h*) How much work does Earth do on the block? (*i*) Draw energy diagrams for three systems during the lifting: block; block and Earth; block, Earth, pulleys, rope, and person. ●●●

Figure P9.31



9.6 Work done on a many-particle system

32. You push a blob of gelatin with a constant force of 3.0 N across a wet table on which it slides easily. Because the blob shape distorts, its center of mass moves only 30 mm during the time interval in which the point of application of your force moves 50 mm. (*a*) What is the work done by you on the blob? (*b*) By how much does your force change the kinetic energy of the center of mass? ●

33. A 1000-kg car traveling at 5.0 m/s runs into the side of an underpass. The concrete wall is not affected, but the car crumples such that in the time interval from the instant of impact to the instant at which the car comes to rest, its center of mass travels forward 0.50 m. What are (*a*) the magnitude of the average force exerted on the car, (*b*) the work done by the wall on the car, and (*c*) the change in kinetic energy of the system? ●

34. A 60-kg ice skater stands facing a wall with his arms bent and then pushes away from the wall by straightening his arms. At the instant at which his fingers lose contact with the wall, his center of mass has moved 0.50 m, and at this instant he is traveling at 3.0 m/s. What are (*a*) the average force exerted by the wall on him, (*b*) the work done by the wall on him, and (*c*) the change in the kinetic energy of his center of mass? ●

35. Two 0.25-kg carts, one red and one green, are stationary on a low-friction track. The carts are initially separated by 0.50 m. You push on the red one with a constant force of 1.5 N toward the green one. When the carts meet, they join together, and you continue to push with the same constant force as the connected carts move a further 0.75 m along the track, at which point you remove your hand. (*a*) What is the work done by you on the two-cart system? (*b*) How far does the system's center of mass move while you are applying the force? (*c*) By what amount does the force change the kinetic energy of the system? ●●

36. Two 0.50-kg carts are pushed toward each other from starting positions at either end of a 6.0-m low-friction track. Each cart is pushed with a force of 3.0 N, and that force is exerted for a distance of 1.0 m. (*a*) What is the work done on the two-cart system? (*b*) What is the change in kinetic energy of the system? (*c*) What is the kinetic energy of the center of mass of the system? ●●

37. You have a 2.0-m chain lying on the floor alongside ten 0.10-kg cubical blocks, each 0.20 m on a side. Each block is resting on the floor, and the inertia of the chain is 1.0 kg. Which process requires less work: lifting one end of the chain so that the chain hangs vertically with the bottom link just touching the floor, or stacking the blocks in a column ten blocks tall? ••

38. A 1.0-kg cart and a 0.50-kg cart sit at different positions on a low-friction track. You push on the 1.0-kg cart with a constant 2.0-N force for 0.15 m. You then remove your hand, and the cart slides 0.35 m and strikes the 0.50-kg cart. (*a*) What is the work done by you on the two-cart system? (*b*) How far does the system's center of mass move while you are pushing the 1.0-kg cart? (*c*) By what amount does your force change the kinetic energy of the system's center of mass? ••

39. Two 1.0-kg blocks, one gray and one tan, are lined up along a horizontal x axis. The gray block is at $x = -4.0$ m, and the tan one is at $x = +4.0$ m. A constant force of 1.0 N is exerted on the gray block in the positive x direction through a distance of 2.0 m. Taking the two blocks as the system and ignoring friction, calculate the following over the interval required for the gray block to move from $x = -4.0$ m to $x = -2.0$ m: (*a*) the work done on the system, (*b*) the change in the energy of the system, (*c*) the change in the system's translational kinetic energy, and (*d*) the value of $F\Delta x_{cm}$, where $F = 1.0$ N and Δx_{cm} is the displacement of the system's center of mass. (*e*) If the blocks collide totally inelastically, what is the energy of the system after the collision? ••

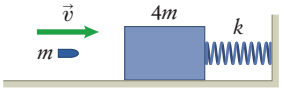
40. Two identical 0.50-kg carts, each 0.10 m long, are at rest on a low-friction track and are connected by a spring that is initially at its relaxed length of 0.50 m and is of negligible inertia. You give the cart on the left a push to the right (that is, toward the other cart), exerting a constant 5.0-N force. You stop pushing at the instant when the cart has moved 0.40 m. At this instant, the relative velocity of the two carts is zero and the spring is compressed to a length of 0.30 m. A locking mechanism keeps the spring compressed, and the two carts continue moving to the right. (*a*) What is the work done by you on the two-cart system? (*b*) How far does the system's center of mass move while you are pushing the left cart? (*c*) By what amount do you change the system's kinetic energy? ••

41. Because the soles of your shoes have cleats, you can exert a backward force of 100 N on ice, allowing you to exert a similar forward force if you throw an object. A 10-kg picnic cooler is at rest on a frozen pond, and you want to get it to shore, up the bank, and stuck in the snow. From past experience, you know that, in order to move up the bank and stick, the cooler needs to be moving at 3.0 m/s at the instant it starts up the bank. Standing 1.0 m from the cooler and not wanting to walk to it, you decide to apply a constant force to throw 1.0-kg snowballs at it to get it moving. If each snowball smashes into the cooler in a totally inelastic collision, what minimum number of snowballs must you throw? •••

42. A bullet of inertia m traveling at speed v is fired into a wooden block that has inertia $4m$ and rests on a level surface. The bullet passes through the block and emerges with speed $v/3$, taking a negligible amount of the wood with it. The block moves to the right but comes to rest after traveling a distance d. (*a*) What is the magnitude of the frictional force between the block and the surface while the block is moving? (*b*) What is the ratio of the energy dissipated as the bullet passes through the block to the energy dissipated by friction between the surface and the bottom face of the block? •••

43. You fire a bullet of inertia m into the block shown in Figure P9.43. The bullet is initially traveling at speed v, and the inertia of the block is $4m$. The surface on which the block sits is rough, and the spring has spring constant k. The bullet becomes embedded in the block. If the spring compresses a maximum distance d, what is the magnitude of the frictional force exerted by the surface on the block while the block is moving? •••

Figure P9.43



9.7 Variable and distributed forces

44. Spring A is stiffer than spring B. You stretch the springs such that the energy stored in each is the same. (*a*) Which spring has the greater change in length from its relaxed position? (*b*) Which spring requires the greater applied force? •

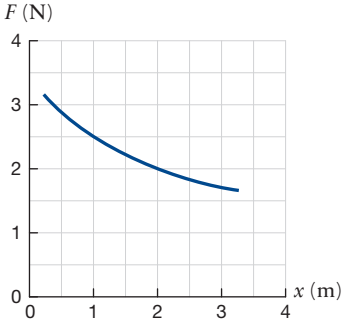
45. Which arrangement in Figure P9.45 requires more work to give the block a rightward displacement $\vec{d}$? The springs, blocks, and surfaces are identical in the two arrangements and the springs are initially at their relaxed length. •

Figure P9.45



46. Figure P9.46 shows the force $F(x)$ plotted as a function of position x. (*a*) What average constant force applied between $x = 1$ m and $x = 3$ m will do the same amount of work? (*b*) What is the work done by $F(x)$ as the object moves from $x = 3$ m to $x = 1$ m? •

Figure P9.46

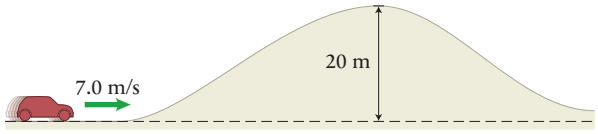


47. Compressing a certain spring to 50% of its relaxed length requires 28 J of work. How much work would be required to compress the same spring by only half as much, to 75% of its relaxed length? •
48. A force varies with time according to the expression $F = a\Delta t$, where $a = 2.0$ N/s. From this information, can you determine the work done on a particle that experienced this force over a displacement of 0.50 m? ••
49. In pushing a 0.024-kg dart into a toy dart gun, you have to exert an increasing force that tops out at 6.0 N when the spring is compressed to a maximum value of 0.12 m. (*a*) What is the launch speed of the dart when fired horizontally? (*b*) Does your answer change if the dart is fired vertically? ••
50. The work done by a certain force is given by $W(\Delta x) = a\Delta x + b(\Delta x)^3$. Write an expression for the force as a function of Δx. ••
51. A 6.0-kg bowling ball is held 10 mm above a mattress and then released from rest and allowed to fall, sinking into the mattress. Model the mattress as a single spring with spring constant $k = 500$ N/m. (*a*) Calculate the ball's kinetic energy, the gravitational potential energy of the system (ball, spring, and Earth), and the elastic potential energy at the following compressions: 0, 0.050 m, 0.10 m, 0.15 m, 0.20 m. (*b*) What is the maximum compression of the mattress? Ignore the inertia of the spring. ••
52. In a carnival game, the player throws a ball at a haystack. For a typical throw, the ball leaves the hay with a speed exactly one-half of the entry speed. (*a*) If the frictional force exerted by the hay is a constant 6.0 N and the haystack is 1.2 m thick, derive an expression for the typical entry speed as a function of the inertia of the ball. Assume horizontal motion only, and ignore any effects due to gravity. (*b*) What is the typical entry speed if the ball has an inertia of 0.50 kg? ••
53. A proton has mass m_p and charge $+e$. The force it exerts on another proton at a distance r is given by $F(r) = \frac{1}{4\pi\epsilon_0}\frac{e^2}{r^2}$ where $\frac{1}{4\pi\epsilon_0} \approx 9 \times 10^9 \text{N} \cdot \text{m}^2/\text{C}^2$, and $e = 1.6 \times 10^{-19}$ C. How much work is needed to increase the separation between the protons from R to $9R$? ••
54. You devise a wind-up car powered by a spring for trips to the grocery store. The car has an inertia of 500 kg and is 4.2 m long. It should be able to accelerate from rest to 20 m/s at least 50 times before the spring needs winding. The spring runs the length of the car, and a full winding compresses it to half this length. In order to meet the acceleration requirement, what must the value of the spring constant be? ••
55. A 0.15-kg cart on wheels is at rest at the origin of an x axis between (but not connected to) two relaxed springs aligned along that axis. When the cart is pushed to the left of the origin, the left spring ($k_l = 4.0$ N/m) is compressed and exerts a force on the cart. When the cart is pushed to the right of the origin, the right spring ($k_r = 6.0$ N/m) is compressed and exerts a force on the cart. (*a*) Plot the force exerted on the cart as a function of its displacement. (*b*) If the cart is given an initial speed of 4.0 m/s at the origin $x = 0$, it might be headed to the left or to the right. How far does it go in each case? (*c*) What is the work done on the cart as it moves from $x = 0$ to its extreme right position and from $x = 0$ to its extreme left position? Ignore any friction effects. •••
56. To complement the baseball-pitching machines at the sports complex where you work, you have designed a catching machine using a mitt, a spring, and a latch that locks the spring in place when the ball has stopped moving. Your boss wants to know the speed of the baseballs coming out of a pitching machine that's been breaking down lately, and you figure you can use the compression of the spring in your catching machine to determine this speed. The mitt that catches the balls has three times the inertia of one ball. Derive an expression for the speed of a thrown ball in terms of the ball's inertia, the distance the spring is compressed, and the spring constant. Take the catch to be a totally inelastic collision between mitt and ball. •••

9.8 Power

57. A particular cyclist can produce an average power output of 180 W when cycling on level ground. While doing this, their speed is 9.5 m/s. What is the magnitude of the net external dissipative force (predominantly air resistance) acting on the cyclist and their bike? •
58. What is the average power delivered when you raise a 2-kg object by 50 cm in 50 ms? •
59. Hiking trails on steep slopes often zigzag back and forth rather than running in a straight path up the slope. What does having a zigzag path accomplish, given that getting to the top of the slope requires the same amount of energy regardless of the path? •
60. The 1000-kg car at the bottom of the hill in Figure P9.60 wants to make it to the top of the hill. What should be the minimum power output of the engine given that 30% of the engine's power and 30% of the car's kinetic energy are lost, and that it takes the car 20 min to make it to the top? ••

Figure P9.60



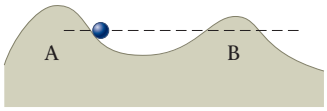
61. A constant external force F is the only force acting on an object. What is the work done on the object by F as a function of time? ••
62. A team of dogs accelerates a 200-kg dogsled from 0 to 5.0 m/s in 3.0 s. (*a*) What is the magnitude of the force exerted by the dogs on the sled? (*b*) What is the work done by the dogs on the sled in the 3.0 s? (*c*) What is the instantaneous power of the dogs at the end of the 3.0 s? (*d*) What is their instantaneous power at 1.5 s? Assume that the acceleration is constant. ••
63. A 20-kg child wants to slide along a horizontal yard toy that consists of a slick plastic sheet covered with a thin layer of water to reduce friction to almost zero. To cause him to slide, you squirt him with a stream of water from a hose. (*a*) If the water from the hose exerts a constant force of 5.0 N on the child, what is the work done by the water in the interval from $t = 0$ to $t = 1.0$ s? (*b*) What is the instantaneous power delivered by the water at $t = 1.0$ s? ••

64. A cog system on the beginning segment of a roller coaster needs to get 25 occupied cars up a 100-m vertical rise over a time interval of 60 s. Each car experiences a gravitational force of 5670 N. The cars start at rest and end up moving at 0.50 m/s. (*a*) How much work is done by the cog system on the cars? (*b*) What average power must the cog system supply? (*c*) What would happen to the amount of work done by the cog system if the cars had to be lifted the same distance in 30 s? ●●
65. A box slides across a frozen pond toward the left shore as two children standing on opposite shores pull on the box with ropes. The child on the left shore pulls with a force of 3.0 N, and the child on the right shore pulls with a force of 2.0 N. At a certain instant, the speed of the box is 3.0 m/s. (*a*) At that instant, what is the power resulting from each force? (*b*) What is the power resulting from both forces? (*c*) Do your answers to parts *a* and *b* change as the box slides? ●●
66. A motor must lift a 1000-kg elevator cab. The cab's maximum occupant capacity is 400 kg, and its constant "cruising" speed is 1.5 m/s. The design criterion is that the cab must achieve this speed within 2.0 s at constant acceleration, beginning from rest. (*a*) When the cab is carrying its maximum capacity, at what rate must the motor deliver energy to get the cab up to cruising speed? (*b*) At what constant rate must the motor supply energy as the fully loaded cab rises after attaining cruising speed? ●●●

Additional Problems

67. A box of inertia 10 kg is suspended by a vertical rope and lowered through a distance of 10 m with a constant speed. How much work is done on the box by the gravitational force? How much work is done by the tension in the rope? ●
68. Is it possible for a force to be exerted on an object *along the direction of the object's motion* and still produce no change in the object's kinetic energy? If yes, give an example. If no, explain why not. ●
69. A shopping cart that has an inertia of 12 kg when empty is loaded with 38 kg of groceries. A child pushing the loaded cart loses control, and the cart rolls into a concrete lamppost, which sustains no damage. The cart is moving at 2.0 m/s when it strikes the lamppost, and the groceries all crash forward on impact, moving the center of mass forward by 0.10 m before everything comes to a halt. What are (*a*) the average force exerted by the post on the cart-groceries system, (*b*) the work done by the post on the system, and (*c*) the change in kinetic energy of the system's center of mass? ●
70. An object is moving in a circle with constant speed. Its kinetic energy is constant since the speed does not change. Would you conclude that the total external force acting on it is zero? ●●
71. To be rescued from the roof of a burning building, a 70-kg man must be able to jump straight up and catch hold of a rope ladder hanging from a helicopter. He has to raise his center of mass 2.00 m above the roof to do so. (*a*) If he crouches down such that his center of mass is 0.50 m above the roof and his feet leave the roof when his center of mass is 1.05 m above the roof, what minimum force must he exert on the roof to reach the ladder? (*b*) A typical candy bar contains about 850 kJ of food energy. How many candy bars does this jump require? ●●
72. About how much work does a trash collector do in the course of an 8-hour shift? ●●
73. A constant, nondissipative external force is applied to a solid block, and it is found that the rate at which energy is delivered by the force is constant as the block moves horizontally. What does this imply for the motion of the block? ●●
74. An object is suspended from the ceiling by a spring aligned along a vertical x axis, with upward defined as the positive x direction. When the spring is at its relaxed length, the object is at the origin of the axis. The object is pulled down to position $-x_1$, held at rest, and released. Consider its motion from the instant of release until the instant it returns to $-x_1$. Ignore friction, air drag, and the inertia of the spring. (*a*) What force(s) is exerted on the object during the round trip, and in what direction(s)? (*b*) How much work is done on the object by the spring as the object travels away from $-x_1$ and back to $-x_1$? ●●
75. You are midway up hill A in Figure P9.75. Across the valley is the shorter hill B, the summit of which is slightly higher than your present spot. You want to roll a ball so that it crests hill B. To get the ball moving, you can roll it either up or down hill A, giving the ball the same initial speed in both cases. (*a*) Which roll direction is more likely to get the ball to crest B, or does it matter? Ignore any energy dissipation. (*b*) In what way would your answer change if you take energy dissipation into account? ●●

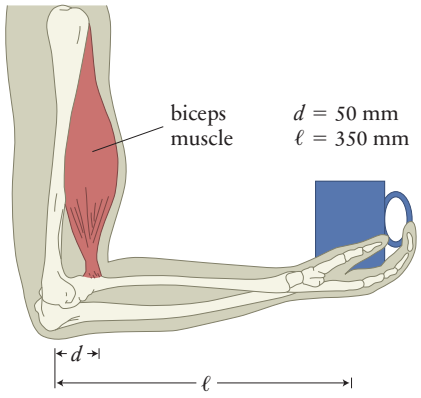
Figure P9.75



76. A 10-kg dog jumps up in the air to catch a ball. The dog's center of mass is normally 0.20 m above the ground, and he is 0.50 m long. The lowest he can get his center of mass is 0.10 m above the ground, and the highest he can get it before he can no longer push against the ground is 0.30 m. If the maximum force the dog can exert on the ground in pushing off is 2.5 times the gravitational force Earth exerts on him, how high can he jump? ●●
77. Assume that when you stretch your torso vertically as much as you can, your center of mass is 1.0 m above the floor. The maximum force you can exert on the floor in pushing off is 2.3 times the gravitational force Earth exerts on you. How low do you have to crouch in order to jump straight up and have your center of mass be 2.0 m above the floor? Is this crouch practical? ●●
78. The 2000-kg cab of a freight elevator is rated to carry a maximum load of 1200 kg, but a careless worker fills the cab with freight that has a combined inertia of 1400 kg. The worker then raises the cab 3.1 m at a steady 2.0 m/s. You notice an unfamiliar whine and come over to investigate. As supervisor, you know that the motor has a 10% tolerance in its rated power output, and you must decide whether there's going to be trouble. ●●● CR
79. Taking a sip of coffee one day, you start thinking about the effort it takes to do so. Your forearm is a type of lever. As shown in Figure P9.79, the biceps muscle is attached to the forearm bone about 50 mm in front of

the elbow joint. The cup of coffee in your hand is 350 mm from the elbow, and your forearm is essentially horizontal. The biceps does work on the forearm as you raise the cup very slowly. You realize that you can use energy principles to estimate the forces involved. ●●● CR

Figure P9.79



80. You are planning to jump off a bridge with a bungee cord tied to your ankles. The bridge deck is 150 m above the water, and the cord spring constant is 40 N/m. Because you must fall the length of the unstretched cord before it begins to stretch, you realize that the unstretched length of the bungee cord, which is adjustable, has to be adjusted based on your inertia. You enjoy thrills and want to reverse your motion just above the water surface. ●●● CR



10

Motion in a Plane

CONCEPTS

10.1 *Straight* is a relative term
10.2 Vectors in a plane
10.3 Decomposition of forces
10.4 Friction
10.5 Work and friction

QUANTITATIVE TOOLS

10.6 Vector algebra
10.7 Projectile motion in two dimensions
10.8 Collisions and momentum in two dimensions
10.9 Work as the product of two vectors
10.10 Coefficients of friction

So far, we have restricted our study of physics to events that take place along a straight line—that is, in one dimension. In this chapter we develop the tools that allow us to deal with motion in a plane—in other words, motion that takes place in two dimensions. As you will see, any problem in two dimensions can be reduced to two one-dimensional problems (which you already know how to solve). After discussing the techniques that allow us to deal with vector quantities in two dimensions, we study the decomposition of two-dimensional force vectors. This decomposition allows us to examine more closely the contact forces that arise at the surface between two objects, which leads us to the subject of friction. Finally, we extend the concept of work so that it applies to situations where the force and force displacement vectors are not along the same line.

10.1 *Straight* is a relative term

To begin our discussion of motion in a plane, look at the video clip in **Figure 10.1**, which shows the motion of a ball dropped from the top of a pole attached to a cart that is moving at constant velocity along a horizontal track. You already know the resulting motion of the ball: It accelerates downward because of the force of gravity exerted on it (see Section 8.8). Indeed, the ball falls alongside the pole. As we discussed in Chapter 6, this is the way the motion must be because the laws of physics are the same in all inertial reference frames, regardless of whether the reference frames are at rest or moving at constant velocity with respect to the surface of Earth.

To extend this example to a more common situation, consider what happens when you drop an object inside the cabin of an airplane cruising at constant velocity. That object falls to the cabin floor in a straight line, exactly as it would if the airplane were at rest on the ground. Likewise, in the reference frame of the moving cart of Figure 10.1, the ball drops to the moving cart with a velocity that points vertically downward. In the cart's reference frame, the motion of the ball is free-fall motion, and a camera that moves along with the cart registers the motion shown in **Figure 10.2*a*.**

The picture is different, however, when observed from a stationary point located somewhere outside the cart—that is, from the Earth reference frame. Because the cart is moving to the right, the trajectory of the ball cannot be a straight vertical line in this reference frame. Figure 10.2*b* shows all the frames of the video clip of Figure 10.1 superimposed on one another. As you can see, in addition to its straight-downward motion relative to the cart, the ball is displaced horizontally from one frame to the next. Note that in spite of the rightward motion of the cart, the distance between ball and pole does not change. This constant distance is precisely what gives the illusion that the ball falls straight down! Your eyes tend to move along with the cart, and

Figure 10.1 A ball dropped from a pole attached to a cart moving at constant velocity falls straight down *relative to the cart*.

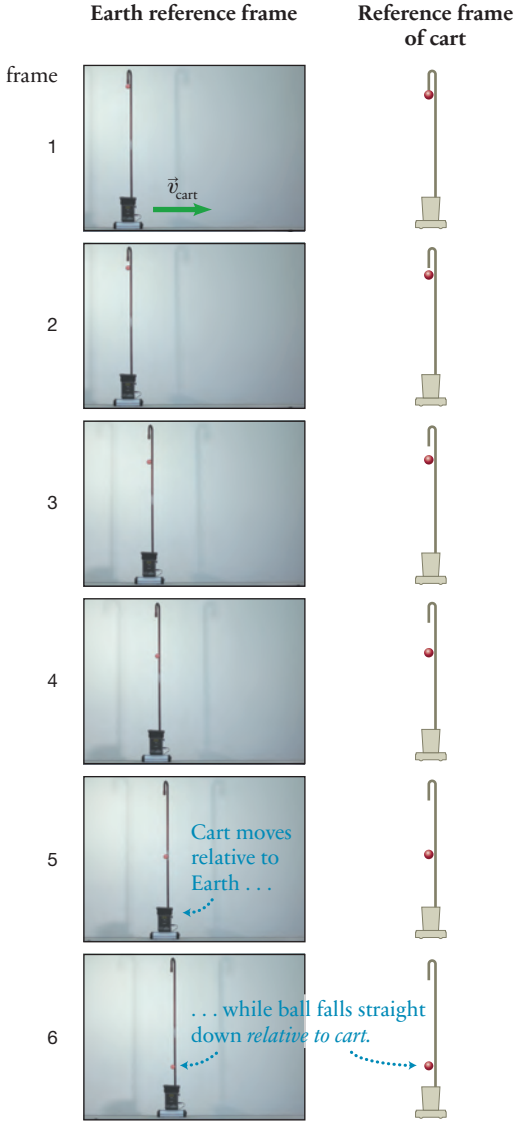
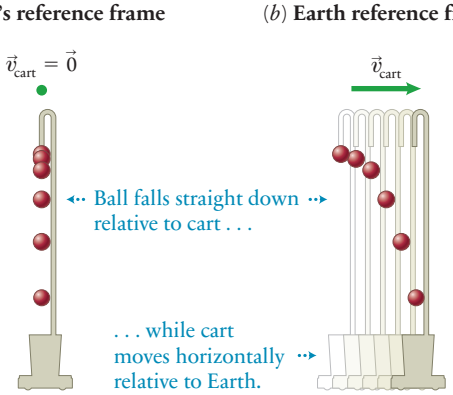


Figure 10.2 Motion of the ball of Figure 10.1 as recorded by a camera that is (*a*) moving along with the cart and (*b*) stationary.



CONCEPTS

so you register only the vertical motion of the ball relative to the cart. The frames of the video clip, however, show that the horizontal displacement of the ball is identical to that of the cart: If the x component of the displacement of the cart is Δx from one frame to the next, then the x component of the displacement of the ball must also be Δx. So, we see that *straight* is a relative term, one that depends on the reference frame.

The analysis in Figure 10.2 tells us that the motion of the ball in the Earth reference frame can be broken down into two parts: free fall in the vertical direction (called the *vertical component* of the motion) and motion at constant velocity in the horizontal direction (the *horizontal component* of the motion). You already know how to treat each of these motions separately. All that is left to do, therefore, is to develop a way to describe both motions at the same time.

10.1 (*a*) In Figure 10.2, what is the ball's velocity the instant before it is released? (*b*) Is the ball's speed in the reference frame of the cart greater than, equal to, or smaller than its speed in the Earth reference frame?

10.2 Vectors in a plane

To describe the motion of the ball in Figure 10.1, we need to generalize our definition of vectors. First of all, we need not one but two reference axes—one for each component of the motion. **Figure 10.3** shows one possible choice for a set of perpendicular reference axes. To distinguish the two axes, we label the horizontal one x and the vertical one y. The origins of both axes coincide with the point at which the ball is released.

To specify the position of the ball in each frame, we need two coordinates: an x *coordinate* to specify position along the x axis and a y *coordinate* to specify position along the y axis. For example, in the coordinate system shown in Figure 10.3, the position of the ball in frame 6 of Figure 10.1 is specified by the coordinates $x = +0.22$ m, $y = -0.42$ m. These coordinates also determine the position vector of the ball (see Section 2.6), which points from the origin to the position of the ball, as shown in Figure 10.3.

Figure 10.3 Position vector for the ball in frame 6 of Figure 10.1, using reference axes in the Earth reference frame.

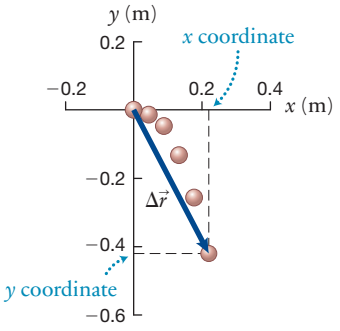
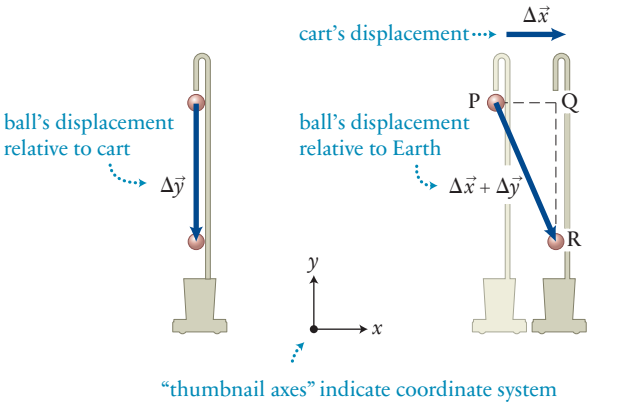


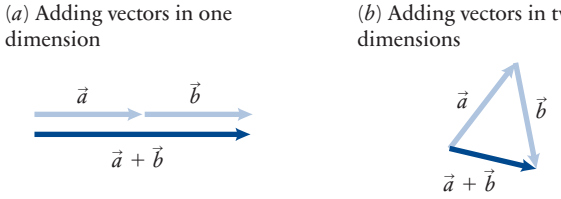
Figure 10.4 Displacement of the ball of Figure 10.1 in the reference frames of the cart and Earth.



Unlike vectors in one dimension (describing motion along a straight line), vectors in two dimensions (describing motion in a plane) do not necessarily point along the axes. Consequently, we need to reexamine the rules for vector addition and subtraction. We'll do this by looking at the vector that represents the ball's displacement between frames 2 and 6 in Figure 10.1. **Figure 10.4*a*** shows the ball's displacement in the reference frame of the moving cart. This vector is labeled $\Delta\vec{y}$ to reflect the fact that it is parallel to the y axis. Figure 10.4*b* shows the corresponding view in the Earth reference frame. The *cart's* displacement is labeled $\Delta\vec{x}$ because it is parallel to the x axis. The *ball's* displacement, however, now points diagonally from the ball's initial position P to its final position R. Note that the ball can also go from P to R by two successive displacements: first from P to Q and then from Q to R. Even if the ball follows this path, though, its displacement is still given by the vector pointing from P to R. In other words, the ball's displacement in the Earth reference frame is the *vector sum* of the horizontal displacement $\Delta\vec{x}$ and the vertical displacement $\Delta\vec{y}$.

As with vectors along a straight line (one dimension), the vector sum of two vectors in a plane (two dimensions) is obtained by placing the tail of the second vector at the head of the first vector. There is, however, one important difference between one-dimensional and two-dimensional vector addition (**Figure 10.5**). In one

Figure 10.5 Vector addition in one and two dimensions.



CONCEPTS

Figure 10.6 The order in which vectors are added does not affect the final answer. These two imaginary flights starting at LaGuardia airport consist of the same vectors in different orders; both end up at the same destination of John F. Kennedy airport.

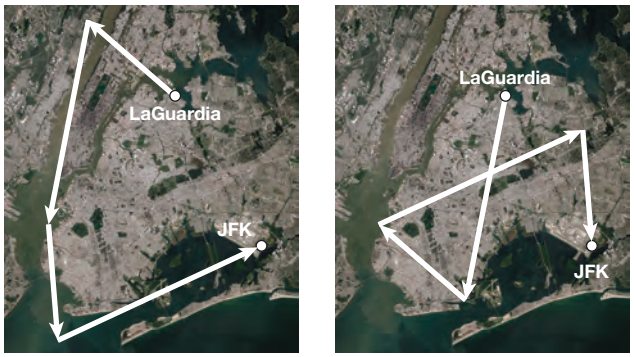
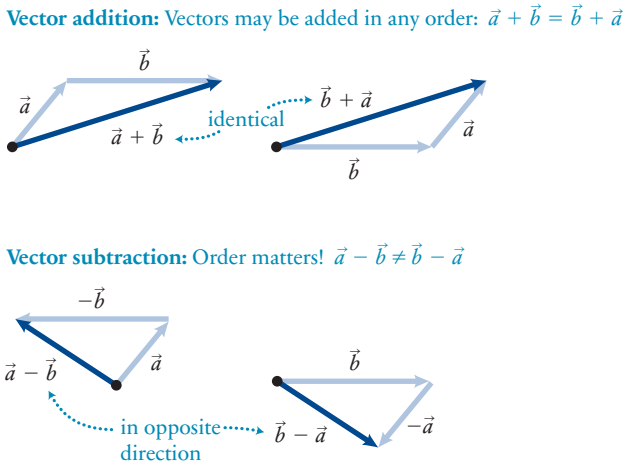


Figure 10.7 Vector addition and subtraction in two dimensions.



dimension, the sum of two vectors of equal length is either zero (when the vectors point in opposite directions) or a vector twice as long as each original vector (when they point in the same direction). In two dimensions, the length of the vector sum depends on the angle between the vectors being added. As Figure 10.5*b* shows, the length of the sum of two vectors can be smaller than the length of either original vector. When two or more vectors in a plane are added, the order in which the vectors are added is immaterial, as **Figure 10.6** shows.

The procedure for subtracting vectors in a plane is identical to that for subtraction in one dimension (**Figure 10.7**):

> **To subtract a vector $\vec{b}$ from a vector $\vec{a}$, reverse the direction of $\vec{b}$ and then add the reversed $\vec{b}$ to $\vec{a}$.**

As in one dimension, the vector $\vec{a} - \vec{b}$ points from the tail of vector $\vec{a}$ to the tip of vector $-\vec{b}$. Note that vector addition is commutative but vector subtraction is not: $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ but $\vec{a} - \vec{b} \neq \vec{b} - \vec{a}$.

We can use the principles of vector addition to develop an important procedure: the decomposition of a vector into components. Just as we can add any two vectors to form a new vector, we can decompose (break up) any vector into two component vectors. For example, any vector $\vec{A}$ can be decomposed into **component vectors** $\vec{A}_x$ and $\vec{A}_y$ along the axes of some conveniently chosen set of mutually perpendicular axes, called a **rectangular coordinate system** (also called a *Cartesian coordinate system*).*

The procedure for decomposing a vector is shown in **Figure 10.8**. Note that if you add the resulting component vectors, you recover the original vector. When you decompose a vector, it is important to cross out the original vector so that you do not take both the original vector and the component vectors into account in any subsequent analysis.

*There is no formal requirement that the axes of the coordinate system be perpendicular to each other. Any set of two intersecting axes allows you to decompose and therefore uniquely specify all possible vectors in a plane. Calculations are considerably simpler, however, with axes that are perpendicular to each other.

Figure 10.8 Procedure for decomposing a vector using a set of mutually perpendicular axes.

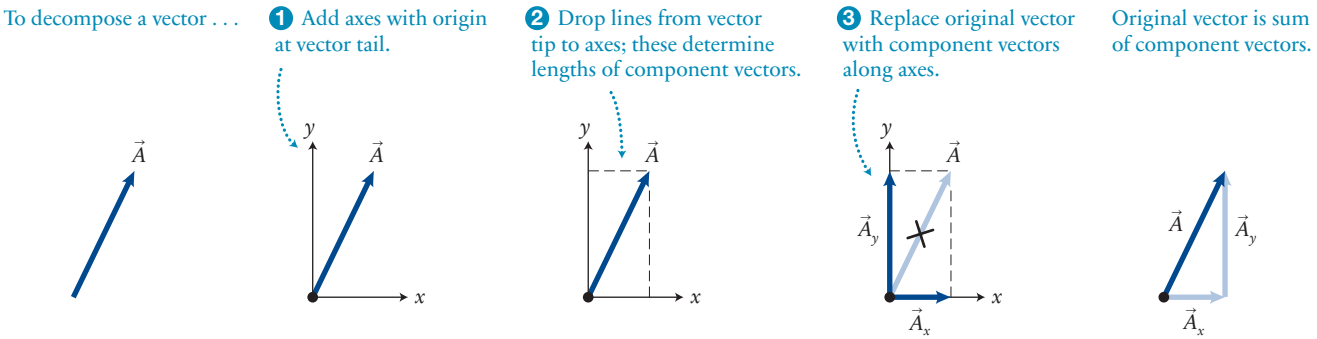


Figure 10.9 Decomposition of a vector $\vec{A}$ in three coordinate systems. In each coordinate system the component vectors $\vec{A}_x$ and $\vec{A}_y$ completely specify the position of point P relative to the origin. Which coordinate system is most convenient depends on the problem at hand.

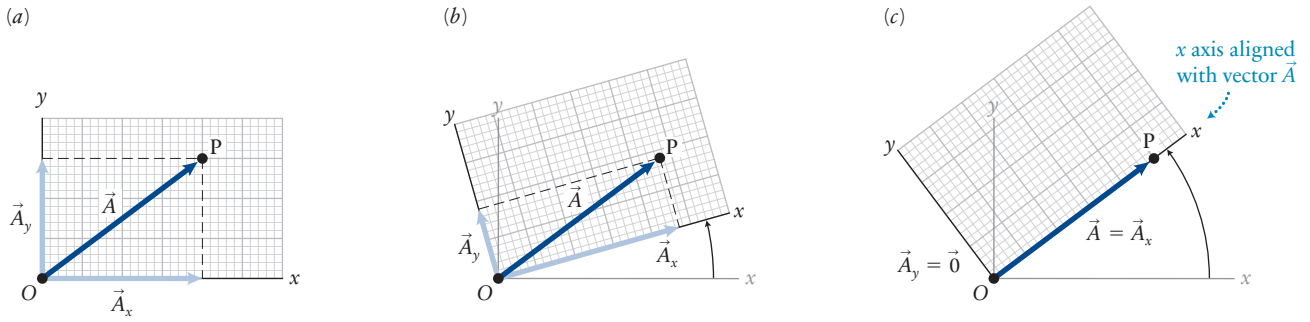


Figure 10.9 shows the decomposition of a vector $\vec{A}$ in three coordinate systems. In each coordinate system the component vectors $\vec{A}_x$ and $\vec{A}_y$ completely specify the position of point P relative to the origin O. In Figure 10.9*c* the coordinate system is such that the vector $\vec{A}$ lines up with the x axis. Now the component vector along the x axis is $\vec{A}$ and the component vector along the y axis is the zero vector. Which choice of coordinate system works best depends on the problem at hand (more on this later).

EXERCISE 10.1 Decompositions

If each major grid unit in Figure 10.9 corresponds to 1 m, specify the location of point P in terms of its x and y coordinates in each of the three coordinate systems.

SOLUTION (*a*) The vertical line dropped from P intercepts the x axis 4 grid units to the right of the y axis, so $x = +4.0$ m. The horizontal line dropped from P gives $y = +3.0$ m. The location of point P is thus given by $(x, y) = (+4.0\text{ m}, +3.0\text{ m})$. ✔

(*b*) $(x, y) = (+4.7\text{ m}, +1.8\text{ m})$. ✔

(*c*) $(x, y) = (+5.0\text{ m}, 0)$. ✔

CONCEPTS

All the vector quantities we have encountered so far obey the rules for addition and subtraction outlined above, and all can be decomposed into component vectors. There is one significant difference, however, between the way vectors are treated in one-dimensional motion and the way they are treated in two-dimensional motion. As we saw in Section 2.8, the displacement and velocity of an object that moves along a straight line point in the same direction. Section 3.1 taught you that if the object speeds up, the acceleration also points in the same direction as the displacement, and if the object slows down, the acceleration points in the direction opposite that of the displacement and velocity. For motion in a plane, the relationship between the directions of these vectors is more complicated.

The displacement, instantaneous velocity, and acceleration of the ball in Figure 10.1 are shown in **Figure 10.10**. Because the ball's average velocity is its displacement divided by the time interval during which this displacement takes place, the average velocity (not shown in Figure 10.10) points in the same direction as the displacement. The ball's instantaneous velocity is the average velocity evaluated over infinitesimally small time intervals, and so if the object does not move in a straight line, the instantaneous velocity does not necessarily point in the same direction as the displacement. In Figure 10.10, for instance, $\Delta\vec{r}$ points from the initial to the final position over a certain time interval, while $\vec{v}$ is tangent to the trajectory—that is, in the instantaneous direction of travel.

Let us next consider the acceleration. By definition, the ball's acceleration points in the direction of the change in velocity. Because the horizontal component of the ball's velocity is equal to the velocity of the cart and so is constant, there is no acceleration in the horizontal direction. Only the vertical component of the ball's velocity changes, and so the ball's acceleration points downward rather than in the direction of $\vec{v}$. If you recall the relationship between acceleration and force ($\sum\vec{F} = m\vec{a}$), this makes sense: The ball's acceleration is caused by the downward force of gravity exerted by Earth on the ball. Therefore both the force and the acceleration point straight down.

What does it mean when the instantaneous velocity and acceleration of an object point in different directions?

Figure 10.10 Displacement, instantaneous velocity, and acceleration for the freely falling ball of Figure 10.1.

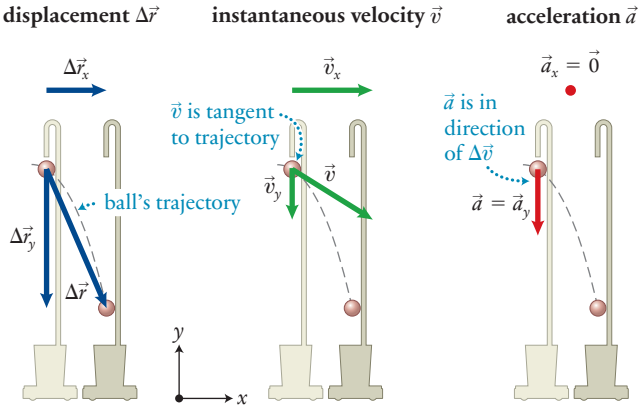
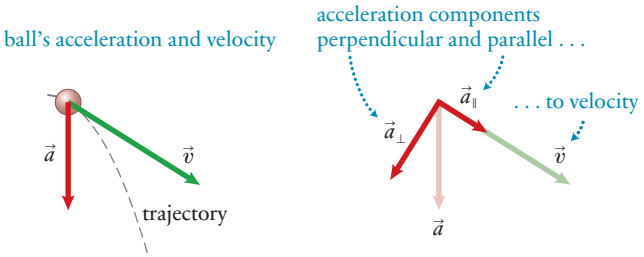


Figure 10.11 Velocity and acceleration vectors of the falling ball in frame 2 of Figure 10.1. Decomposition of $\vec{A}$ into component vectors parallel and perpendicular to $\vec{v}$.



To answer this question, I must decompose the acceleration into two components: one parallel to the instantaneous velocity and one perpendicular to it (**Figure 10.11**). The parallel component causes the length of the instantaneous velocity—in other words, the ball's *speed*—to change. Just as for one-dimensional motion, the speed increases when the parallel component of the acceleration points in the same direction as the instantaneous velocity (as is the case with our ball-and-cart example) and decreases when they point in opposite directions. The perpendicular component of the acceleration causes the *direction* of the instantaneous velocity to change, which in our particular case of the ball and cart means that the ball does not move in a straight horizontal line the way the cart does. Instead, the ball's instantaneous velocity in any given frame of the video clip points more toward the ground than does its counterpart in the preceding frame. (In short, the ball curves downward.) In summary:

> **In two-dimensional motion, the component of the acceleration parallel to the instantaneous velocity changes the speed; the component of the acceleration perpendicular to the instantaneous velocity changes the direction of the instantaneous velocity but not its magnitude.**

These ideas are illustrated in **Figure 10.12**.

Figure 10.12 How the parallel and perpendicular components of acceleration affect velocity.

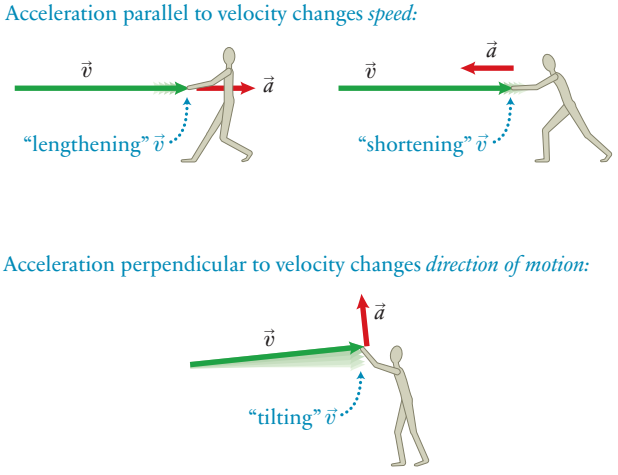
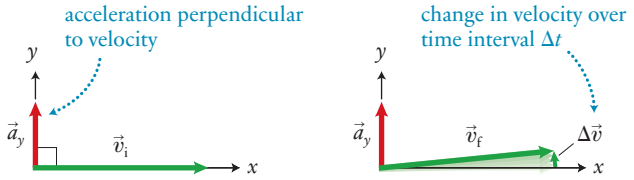


Figure 10.13 Change in velocity caused by a perpendicular acceleration.



If you wonder how it can be that the perpendicular component changes *only* the direction of the velocity and does not affect its magnitude, examine **Figure 10.13**. An object initially moving in the positive x direction is given an acceleration in the positive y direction. Because of this acceleration, over a small but finite time interval Δt the object's initial velocity $\vec{v}_i$ gains a small component $\Delta\vec{v}$ in the y direction. The final velocity $\vec{v}_f = \vec{v}_i + \Delta\vec{v}$ therefore no longer points along the x axis—as we concluded before, $\vec{a}_y$, being perpendicular to $\vec{v}_i$, changes the direction of $\vec{v}_i$. Why then doesn't it also change the magnitude? The Pythagorean theorem tells us that the magnitude of $\vec{v}_f$ must be greater than that of $\vec{v}_i$, a result that appears to contradict what I just said about a perpendicular acceleration not affecting the magnitude of a horizontal velocity. The resolution of this contradiction lies in realizing that, as the direction of $\vec{v}$ changes, so does the angle between $\vec{v}$ and $\vec{a}_y$: The acceleration, because it is no longer exactly perpendicular to the velocity, now has a small component parallel to the velocity, and it is this component of the acceleration that causes the change in the magnitude of $\vec{v}$. We must therefore restrict our analysis to an infinitesimally small time interval. With this restriction, we can ignore the change in the magnitude of the velocity caused by $\Delta\vec{v}$: $v_i \approx v_f$.

CONCEPTS

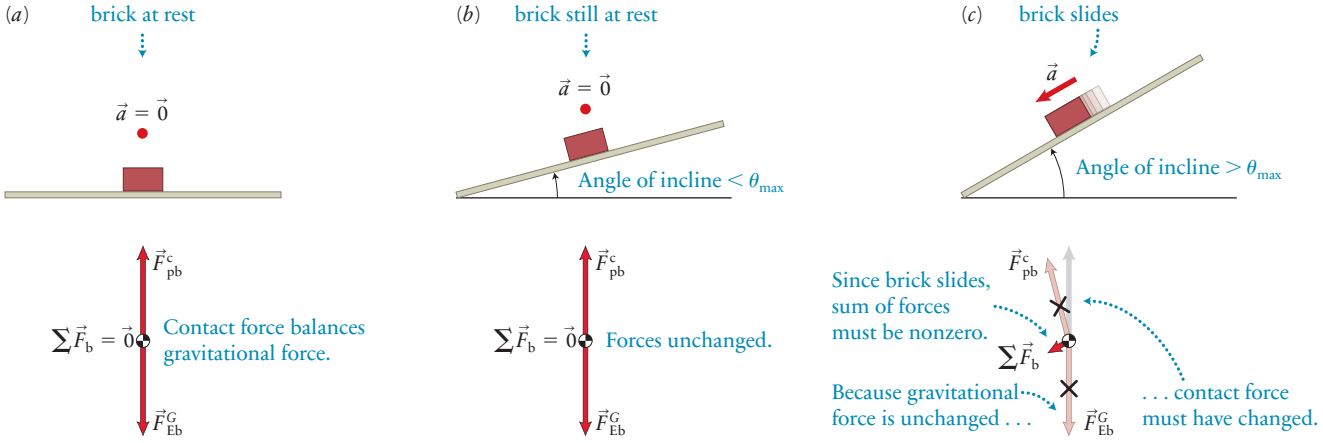
10.2 In Figure 10.10, the ball's instantaneous velocity $\vec{v}$ does not point in the same direction as the displacement $\Delta\vec{r}$ (it points *above* the final position of the ball). Why?

10.3 Decomposition of forces

Now that you know how to decompose vectors into components, let us see how this technique works with forces. Consider, for example, putting a brick on a horizontal plank and then slowly tilting the plank, as illustrated in **Figure 10.14**. Initially, the brick remains at rest, but when the angle of incline θ exceeds some maximum value θ_{max}, the brick suddenly slides downward. The angle of incline at which an object on an inclined surface begins to slide depends on the materials out of which the surface and the object are made and on their smoothness. It is also strongly affected by the presence of dirt or lubricants.

When the plank of Figure 10.14 is horizontal, the forces exerted on the brick are a downward gravitational

Figure 10.14 The forces exerted on a brick as you raise one end of the plank on which it rests.



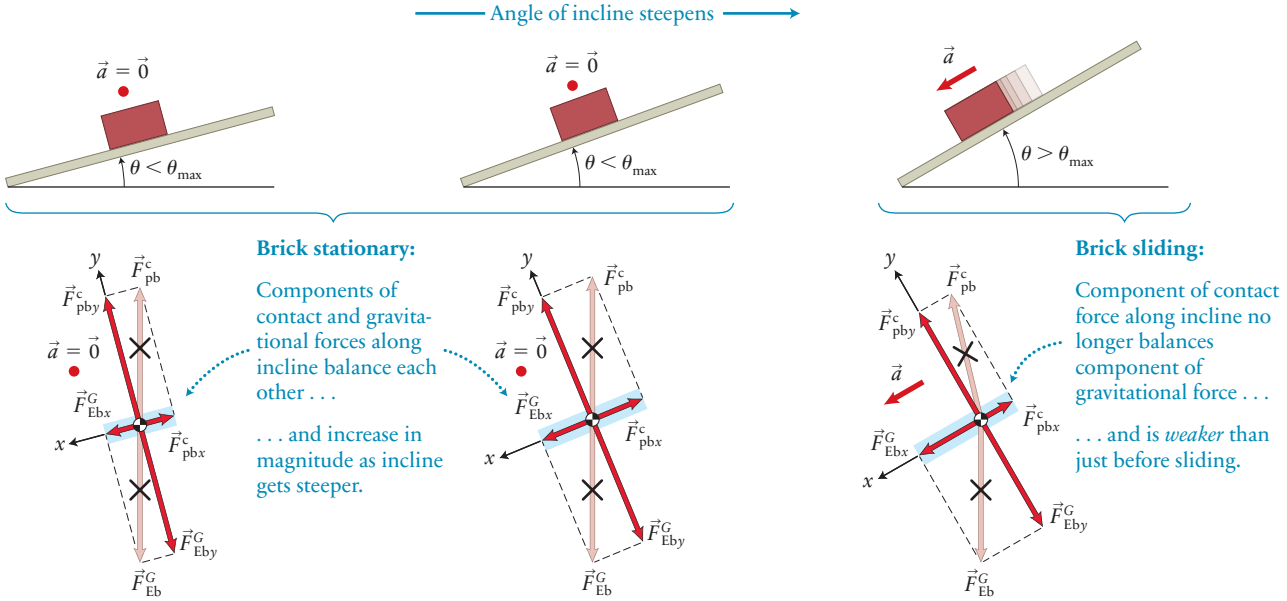
force exerted by Earth and an upward contact force of equal magnitude exerted by the plank, as illustrated in the free-body diagram in Figure 10.14*a*. As long as the brick stays at rest, raising one end of the plank doesn't change the free-body diagram: The force of gravity on the brick does not depend on the inclination of the plank, and because the brick is still at rest we know that the force exerted by the plank must still be upward and equal in magnitude to the force of gravity, as shown in Figure 10.14*b*. When the angle of incline exceeds θ_{max} the situation changes, however. At that point, the contact force provided by the plank is no longer able to balance the downward force of gravity, and the brick accelerates down the incline. Although it is not immediately clear how to draw the free-body diagram, we know that the vector sum of the forces exerted on the brick must point in the same direction as the brick's acceleration (Figure 10.14*c*).

Because the brick is constrained to move along the surface of the plank, its acceleration must always be parallel to the plank, as shown in Figure 10.14*c*. For this reason, it makes sense to analyze the forces in terms of components parallel and perpendicular to the plank, by pointing the x axis along the surface that constrains the motion, as shown in **Figure 10.15**. The force components perpendicular to the surface are called **normal components;** those parallel to the surface are called **tangential components.** Because the brick's acceleration must be parallel to the plank, the normal component vectors $\vec{F}^c_{pby}$ and $\vec{F}^G_{Eby}$ must always be equal in magnitude—whether the brick is at rest or sliding. (If these two forces were not of equal magnitude, the brick would accelerate in the direction perpendicular to the plank, and we know this doesn't happen.)

The tangential component vectors $\vec{F}^c_{pbx}$ and $\vec{F}^G_{Ebx}$, on the other hand, are equal in magnitude only when

CONCEPTS

Figure 10.15 Free-body diagrams for a brick on a plank at different angles of inclination.



the brick is not accelerating down the incline. Note in Figure 10.15 that as the angle of incline is increased, the magnitude of the tangential component vector $\vec{F}^G_{\text{Eb}x}$ of the force of gravity increases. To prevent the brick from sliding down, the magnitude of the tangential component vector $\vec{F}^c_{\text{pb}x}$ of the contact force exerted by the plank on the brick must therefore also increase. When the angle of incline exceeds $\theta_{\max}$, the plank is no longer able to provide the sideways support necessary to hold the brick in place. The tangential component vectors of the forces no longer add up to zero, and the vector sum of the forces exerted on the brick is directed down the plank, as shown in Figure 10.14*c*.

10.3 A suitcase being loaded into an airplane moves at constant velocity on an inclined conveyor belt. Draw a free-body diagram for the suitcase as it moves up along with the belt. Show the normal and tangential components of the forces exerted on the suitcase.

Let us now return to a question raised in Section 10.2: What set of axes should you choose in a given problem? The example of the brick sliding down the inclined plank suggests an answer:

If possible, choose a coordinate system such that one of the axes lies along the direction of the acceleration of the object under consideration.

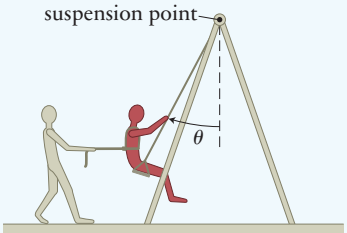
This choice of coordinate system allows you to break the problem neatly into two parts because only force components along the direction of acceleration can contribute to the motion.

When the acceleration of the object being considered is zero, choose your axes to coincide with the directions of as many forces as possible because no decomposition is necessary for any force that lies along an axis.

EXAMPLE 10.2 Pulling a friend on a swing

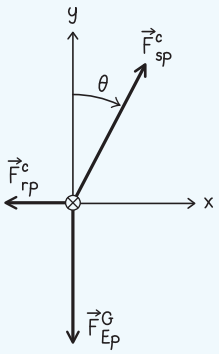
Using a rope, you pull a friend sitting on a swing (**Figure 10.16**). (*a*) As you increase the angle θ, does the magnitude of the force $\vec{F}^c_{\text{rp}}$ required to hold your friend in place increase or decrease? (*b*) Is the magnitude of that force larger than, equal to, or smaller than the magnitude of the gravitational force $\vec{F}^G_{\text{Ep}}$ exerted by Earth on your friend? (Consider the situation for both small and large values of θ.) (*c*) Is the magnitude of the force $\vec{F}^c_{\text{sp}}$ exerted by the swing on your friend larger than, equal to, or smaller than F^G_{Ep}?

Figure 10.16 Example 10.2.



1 GETTING STARTED I begin by drawing a free-body diagram of your friend (**Figure 10.17**). Three forces are exerted on him: $\vec{F}^G_{\text{Ep}}$, the force of gravity directed vertically downward, the horizontal force $\vec{F}^c_{\text{rp}}$ exerted by the rope, and a force $\vec{F}^c_{\text{sp}}$ exerted by the swing seat. This latter force $\vec{F}^c_{\text{sp}}$ is exerted by the suspension point via the chains of the swing and is thus directed along the chains. I therefore choose a horizontal x axis and a vertical y axis, so that two of the three forces lie along axes. Because your friend's acceleration is zero, the vector sum of the forces must be zero.

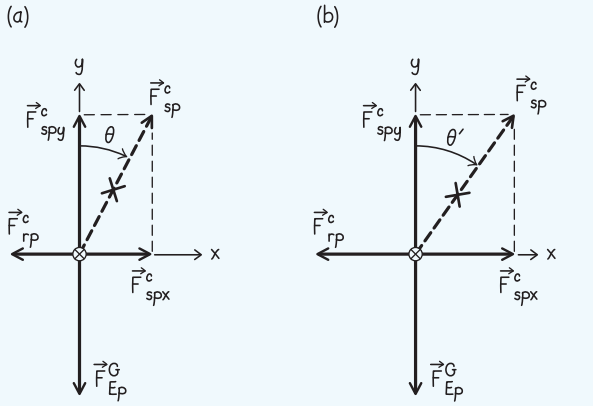
Figure 10.17



2 DEVISE PLAN Because your friend is at rest, the vectors along the two axes must add up to zero. The best way to see how the magnitude of the force $\vec{F}^c_{\text{rp}}$ exerted by the rope must change as θ is increased is to draw free-body diagrams showing different values of θ. To answer parts *b* and *c*, I can compare the various forces in my free-body diagrams.

3 EXECUTE PLAN (*a*) I begin by decomposing $\vec{F}^c_{\text{sp}}$ into x and y components (**Figure 10.18*a***). Because the forces must add to zero along both axes, I conclude from my diagram that $\vec{F}^c_{\text{sp}y}$ must be equal in magnitude to $\vec{F}^G_{\text{Ep}}$, the downward force of gravity. Likewise, $\vec{F}^c_{\text{sp}x}$ must be equal in magnitude to $\vec{F}^c_{\text{rp}}$, the horizontal force the rope exerts on your friend. Next, I draw a second free-body diagram for a larger angle θ (Figure 10.18*b*). As θ increases, $\vec{F}^c_{\text{sp}y}$ must remain equal in magnitude to $\vec{F}^G_{\text{Ep}}$ (otherwise your friend would accelerate vertically). As Figure 10.18*b* shows, increasing θ while keeping $\vec{F}^c_{\text{sp}y}$ constant requires the magnitude of $\vec{F}^c_{\text{sp}x}$ to increase. Because your friend is at rest, the forces in the horizontal direction must add up to zero and so $F^c_{\text{rp}} = |F^c_{\text{sp}x}|$. So if the magnitude of $\vec{F}^c_{\text{sp}x}$ increases, the magnitude of $\vec{F}^c_{\text{rp}}$ must increase, too. ✔

Figure 10.18



(Continued)

CONCEPTS

(b) From Figure 10.18 I see that $\tan\theta = |F^c_{spx}|/|F^c_{spy}|$. For $\theta < 45°$, $\tan\theta < 1$, and so $|F^c_{spx}| < |F^c_{spy}|$. Because $|F^c_{spy}| = F^G_{Ep}$ and $|F^c_{spx}| = F^c_{rp}$, I discover that for $\theta < 45°$, $F^c_{rp} < F^G_{Ep}$. When $\theta > 45°$, $\tan\theta > 1$, and so $|F^c_{spx}| > |F^c_{spy}|$ and $F^c_{rp} > F^G_{Ep}$. ✔

(c) $|\vec{F}^c_{spy}| = F^G_{Ep}$ and $F^c_{sp} = \sqrt{(F^c_{spx})^2 + (F^c_{spy})^2}$. Therefore, F^c_{sp} must always be larger than F^G_{Ep} when $\theta \neq 0$. ✔

4 EVALUATE RESULT I know from experience that you have to pull harder to move a swing farther from its equilibrium position, and so my answer to part a makes sense. With regard to part b, when the swing is at rest at 45°, the forces $\vec{F}^c_{rp}$ and $\vec{F}^G_{Ep}$ on your friend make the same angle with the force $\vec{F}^c_{sp}$, and so $\vec{F}^c_{rp}$ and $\vec{F}^G_{Ep}$ should be equal in magnitude. The force of gravity is independent of the angle, but the force exerted by the rope increases with increasing angle, and so it makes sense that for angles larger than 45°, $\vec{F}^c_{rp}$ is larger than $\vec{F}^G_{Ep}$. In part c, because the vertical component of the force $\vec{F}^c_{sp}$ exerted by the seat on your friend always has to be equal to the force of gravity, adding a horizontal component makes $\vec{F}^c_{sp}$ larger than $\vec{F}^G_{Ep}$, as I found.

10.4 You decide to move a heavy file cabinet by sliding it across the floor. You push against the cabinet, but it doesn't budge. Draw a free-body diagram for it.

10.4 Friction

The force that opposes your push on the file cabinet in Checkpoint 10.4—the tangential component of the contact force exerted by the floor on the cabinet—has to do with friction. If the floor were very slick or if the cabinet had wheels, there would be little friction and your push would easily move the cabinet. Instead, you have to lean against it with all your strength until, with a jerk, it suddenly begins to slide. Once you get the cabinet moving, you must keep pushing to keep it in motion. If you stop pushing, friction stops the motion.

10.5 (a) Suppose you push the file cabinet just enough to keep it moving at constant speed. Draw a free-body diagram for the cabinet while it slides at constant speed. (b) Suddenly you stop pushing. Draw a free-body diagram for the file cabinet at this instant.

Don't skip Checkpoint 10.5! It will be harder to understand the rest of this section if you haven't thought about these situations.

Even though the normal and tangential components of the contact force exerted by the floor on the cabinet belong to the same interaction, they behave differently and are usually treated as two separate forces: the normal component being called the **normal force** and the tangential component being called the **force of friction.**

To understand the difference between normal and frictional forces, consider a brick on a horizontal wooden plank supported at both ends (**Figure 10.19*a***). Because the brick is at rest, the normal force $\vec{F}^c_{pb\perp}$ exerted by the plank on it is equal in magnitude to the gravitational force exerted on it. Now imagine using your hand to push down on the brick with a force $\vec{F}^c_{hb}$. Your downward push increases the total downward force exerted on the brick, and, like a spring under compression, the plank bends until the normal force it exerts on the brick balances the combined downward forces exerted by your hand and by Earth on the brick (Figure 10.19b). As you push down harder, the plank bends more, and the normal force continues to increase (Figure 10.19c) until you exceed the plank's capacity to provide support and it snaps, at which point the normal force suddenly disappears (Figure 10.19d). So, normal forces take on whatever value is required to prevent whatever is pushing down on a surface from moving through that surface—up to the breaking point of the supporting material.

Next imagine that instead of pushing down on the brick of Figure 10.19a, you gently push it to the right, as in **Figure 10.20**. As long as you don't push hard, the brick remains at rest. This tells you that the horizontal forces exerted on the brick add to zero, and so the plank must be exerting on the brick a horizontal frictional force that is equal in magnitude to your push but in the opposite direction. This horizontal force is caused by microscopic bonds between the surfaces in contact. Whenever two objects are placed in contact, such bonds form at the extremities of microscopic bumps on the surfaces of the objects. When you try to slide the surfaces past each other, these tiny bonds prevent sideways motion. As you push the brick to the right, the bumps resist bending and, like microscopic springs, each bump exerts a force to the left. The net effect of all these microscopic forces

CONCEPTS

Figure 10.19 A demonstration of the normal force.

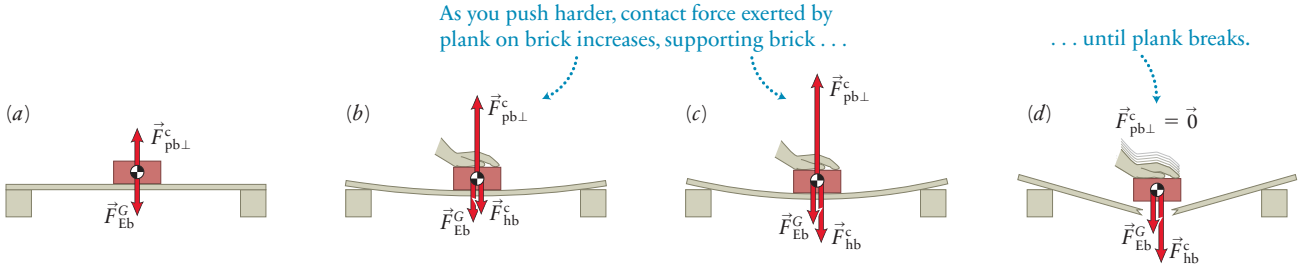
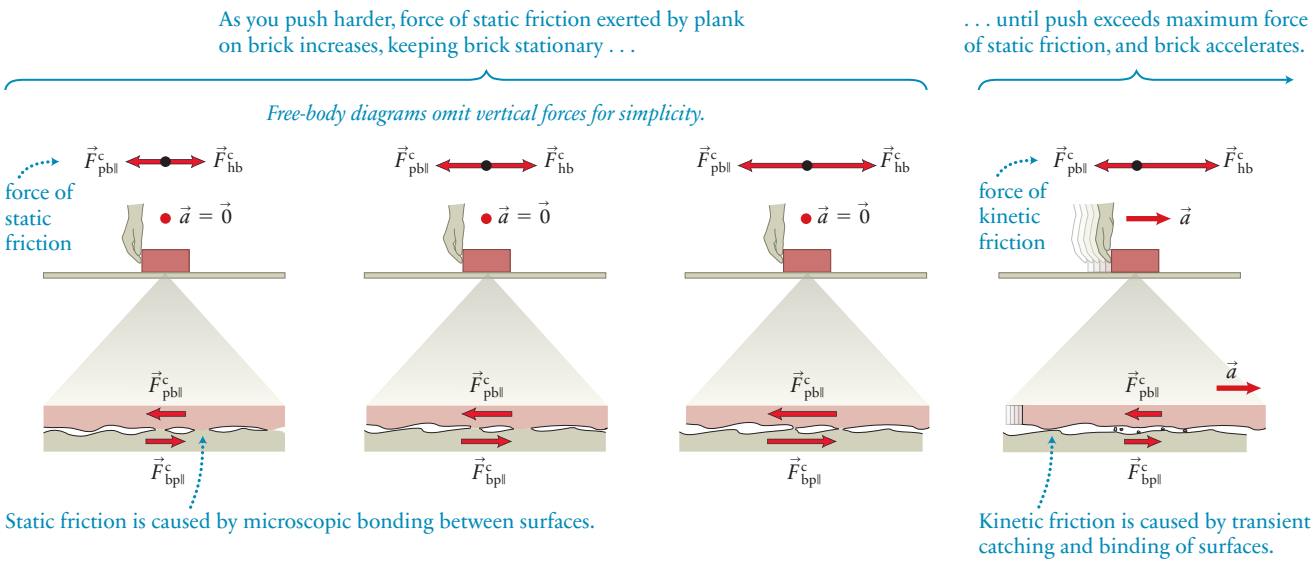


Figure 10.20 The microscopic basis for forces of static and kinetic friction (b = brick, p = plank, h = hand).



is to hold the brick in place. As you increase the force of your push, the bumps resist bending more and the tangential component of the contact force grows. This friction exerted by surfaces that are not moving relative to each other is called **static friction.**

The tiny bonds cannot withstand large forces, however, and as you continue to increase your sideways push on the brick, the bonds get ripped apart, the tangential force they provide can no longer resist the push, and the brick starts to slide. The surface of the brick is no longer at rest relative to the surface on which the brick is resting—the two surfaces are said to *slip*.

Once the brick slips, bumps on the surfaces plow into each other, continually forming and breaking microscopic bonds, and microscopic pieces chip off from the surface. This constant bumping, making and breaking of bonds, and chipping cause what is called **kinetic friction.** Even though kinetic friction causes slipping surfaces to wear and to heat up, it is not always undesirable: Kinetic friction makes it possible to slow down the moving wheels of a car or bicycle, to use an eraser, and to warm up your hands by rubbing them, just to name a few examples.

So, just as the normal force takes on whatever value is necessary to prevent motion on a surface, the force of static friction takes on whatever value is necessary to prevent two surfaces from slipping past each other. The main differences between normal forces and forces of static friction are: (1) The maximum value of the force of static friction is generally much smaller than the maximum value of the normal force (it is relatively easy to push the brick across the plank but nearly impossible to push it through the plank), and (2) once the maximum value of the normal force is reached, the normal force disappears, but once the maximum value of the force of static friction is reached, there still is a smaller but nonzero force of kinetic friction.

To distinguish more easily among the various components of the contact force exerted by surface 1 on surface 2, we shall when necessary replace the superscript c by n for the normal force ($\vec{F}^{c}_{12\perp}$ becomes $\vec{F}^{n}_{12}$), by s for the force of static friction, and by k for the force of kinetic friction ($\vec{F}^{c}_{12\parallel}$ becomes $\vec{F}^{s}_{12}$ or $\vec{F}^{k}_{12}$).

10.6 Which type of friction—static or kinetic—plays a role in (*a*) holding a pencil, (*b*) chalk making marks on a chalkboard, (*c*) skiing downhill, (*d*) polishing a metal surface, and (*e*) walking (without sliding) down an incline?

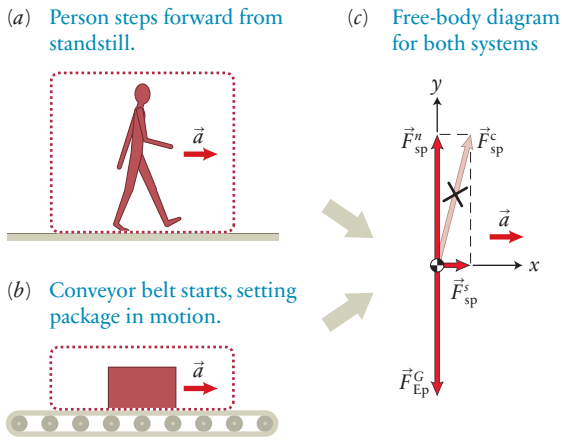
10.5 Work and friction

Like the normal force, the force of static friction is an elastic force: It causes no irreversible changes. The deformations that take place undo themselves when the force that caused them is removed. If you stop pushing on the file cabinet of Checkpoint 10.4 before it starts moving, for example, you will not scratch the floor. Kinetic friction, however, does cause damage, mainly because of the chipping mentioned earlier. Push the cabinet a short distance, no matter how carefully, and you will cause some damage, both to the bottom surface of the cabinet and to the floor surface. At best the damage is microscopic, but if you were to push the cabinet around the floor all day long, your landlady would probably not renew your lease. The heating and wear caused by kinetic friction turn coherent energy into incoherent energy. To sum up:

The force of kinetic friction is not an elastic force and so causes energy dissipation. The force of static friction is an elastic force and so causes no energy dissipation.

CONCEPTS

Figure 10.21 Two equivalent situations in which an object is accelerated by static friction.



In Section 9.4 I cautioned against choosing a system for which friction occurs at the boundary because you don't know how much of the dissipated energy ends up on each side of the boundary. Now you see that this caveat does not apply to static friction because static friction causes no energy dissipation.

This brings us to the question of whether or not the force of static friction can do work on a system. After all, static friction does not involve any motion of the two surfaces relative to each other, and work requires a nonzero force displacement. To answer this question, consider the two situations shown in **Figure 10.21**: A person starts to walk and a package is set in motion by a conveyor belt. In both cases the system (the person in Figure 10.21*a* and the package in Figure 10.21*b*) gains kinetic energy. Where does that energy come from, and what is the role of static friction in each case?

Let us first look at Figure 10.21*a*. To begin walking, the person leans slightly forward and raises one foot, with her other foot pressing on the floor downward and backward. The floor surface then supplies an upward contact force $\vec{F}^c_{sp}$ that is tilted forward, as shown in Figure 10.21*c*. The force $\vec{F}^c_{ps}$ of the person pressing down on the surface and the force $\vec{F}^c_{sp}$ of the surface pushing up form an interaction pair. Because she is not accelerating in the vertical direction, the forces in that direction must add to zero, and so the normal component of the contact force exerted on the person—the normal force $\vec{F}^c_{sp}$—is equal in magnitude to the force of gravity exerted on her. The tangential component $\vec{F}^s_{sp}$ is what propels her forward. As long as the foot doesn't slip, this tangential component is due to static friction; the forward acceleration is then equal to the force of static friction divided by the person's inertia.

To determine whether or not the force of static friction exerted by the floor surface on the person does work on the person, we need to look at the force displacement. Because the force of static friction is applied at the foot, which does not move in any direction while it is on the floor surface, the force displacement is zero. The surface therefore does no work on the person. That makes sense because *she* supplies the energy required to accelerate, not the surface.

Next, look at Figure 10.21*b*. The free-body diagram for this situation is the same as that for the person. Two forces are exerted on the package: a downward force of gravity $\vec{F}^G_{Ep}$ and a contact force $\vec{F}^c_{sp}$ exerted by the surface of the conveyor belt. The normal component of the contact force is equal in magnitude to the force of gravity; the tangential component (the force of static friction) is what causes the package to accelerate. The point of application of the force of static friction moves along with the package, and so the force of static friction does work on the package. This, too, makes sense: The package has no way of internally generating any kinetic energy; it is the conveyor that must do this. So, like any other elastic contact force, the force of static friction *can* do work on a system, and it is perfectly all right to choose a system in which *static* friction occurs at the boundary. It's only kinetic friction that causes problems because this type of friction dissipates energy, generating thermal energy that is distributed on both sides of the surface.

Figure 10.21 also dispels the often-heard myth that friction always *opposes* motion. In both cases shown in the figure, friction *causes* the motion. Without the forward pointing force of friction, neither the person nor the package would accelerate. Friction opposes the motion of two surfaces *relative to each other,* not motion per se. In Figure 10.21*a*, for instance, the person's rear foot would slip backward if there were no forward-pointing force of friction. The force of static friction prevents this slipping motion, and in doing so it pushes her forward.

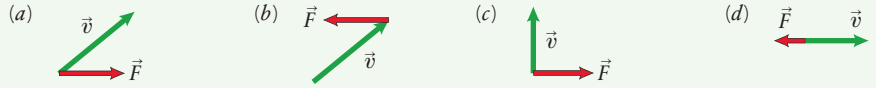
10.7 Draw energy diagrams for the person and the package in Figure 10.21.

CONCEPTS

Self-quiz

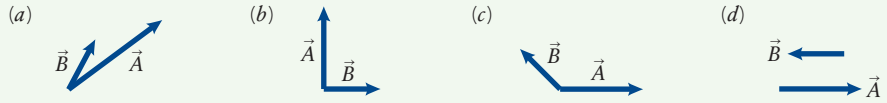
1. In the diagrams in **Figure 10.22**, the velocity of an object is given along with the vector representing a force exerted on the object. For each case, determine whether the object's speed increases, decreases, or remains constant. Also determine whether the object's direction changes in the clockwise direction, changes in the counterclockwise direction, or doesn't change.

Figure 10.22



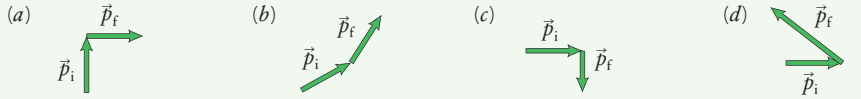
2. Draw the sum $\vec{A} + \vec{B}$ and the difference $\vec{A} - \vec{B}$ of each vector pair in **Figure 10.23**.

Figure 10.23



3. Each diagram in **Figure 10.24** indicates the momentum of an object before and after a single force is exerted on it. For each case determine the direction of the force.

Figure 10.24

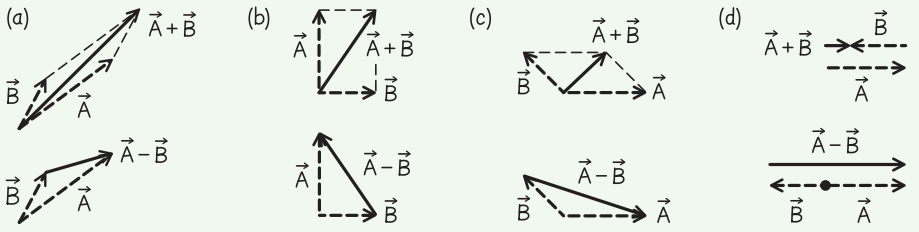


4. While driving to school, you place your coffee cup on the dashboard. Which type of force of friction allows the cup to stay put on the dashboard when the car speeds up? Does this force do any work on the cup?

Answers

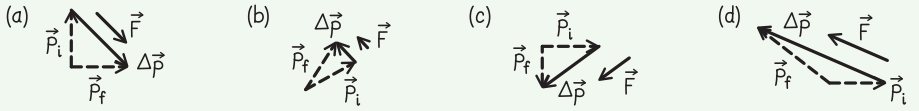
1. (*a*) $\vec{F}$ has component vectors parallel and perpendicular to $\vec{v}$. Because $\vec{F}_{\parallel}$ is in the same direction as $\vec{v}$, the speed increases. $\vec{F}_{\perp}$ tilts $\vec{v}$ clockwise from its current position. (*b*) Because the direction of $\vec{F}_{\parallel}$ is opposite the direction of $\vec{v}$, the speed decreases. $\vec{F}_{\perp}$ tilts $\vec{v}$ counterclockwise from its current position. (*c*) Because $\vec{F}_{\parallel} = 0$, the speed remains the same; $\vec{F}_{\perp}$ tilts $\vec{v}$ clockwise. (*d*) $\vec{F}_{\parallel}$ decreases the speed. $\vec{F}_{\perp} = 0$ and so the direction of $\vec{v}$ does not change.
2. The vector sum is equal to the diagonal of a parallelogram drawn using the two vectors placed tail to tail. The vector difference points from the tip of the second vector to the tip of the first vector (**Figure 10.25**).

Figure 10.25



3. In each case the force is parallel to $\Delta\vec{p} = \vec{p}_f - \vec{p}_i$ (**Figure 10.26**).

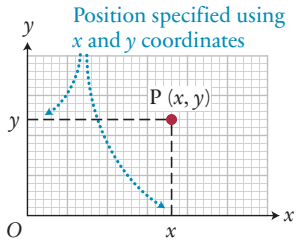
Figure 10.26



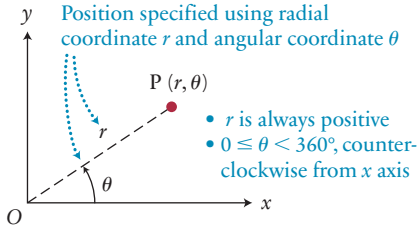
4. Because no sliding occurs, the force of friction must be static. Because the force displacement and the force point in the same direction when the car speeds up, this force does positive work on the cup.

Figure 10.27 Rectangular and polar coordinates.

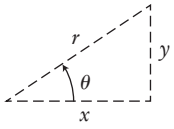
(*a*) Rectangular coordinates



(*b*) Polar coordinates



(*c*) Relationship between the two



10.6 Vector algebra

In a rectangular coordinate system, the position of a point P is specified by the rectangular coordinates x and y (**Figure 10.27a**). Alternatively, we can specify a point using **polar coordinates** (Figure 10.27*b*). The *radial* coordinate r gives the distance from the origin to the point and is always positive. The *angular* coordinate θ specifies the angle between r and the x axis and is measured counterclockwise from the positive x axis ($0 \le \theta < 360°$). Any arbitrary point P is specified entirely by either (x, y) or (r, θ).

As can be seen from Figure 10.27*c*, the radial coordinate r is the hypotenuse of a right triangle with sides x and y, so

$$r = +\sqrt{x^2 + y^2}. \tag{10.1}$$

The plus sign in front of the square root indicates that r is never negative. Trigonometry tells us that $\tan\theta = y/x$, so

$$\theta = \tan^{-1}\left(\frac{y}{x}\right), \tag{10.2}$$

where the signs of x and y determine in which quadrant θ lies. Equations 10.1 and 10.2 allow us to determine polar coordinates from rectangular ones. Using trigonometry, we can also express rectangular coordinates in terms of polar coordinates:

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta. \end{aligned} \tag{10.3}$$

EXERCISE 10.3 Polar point

If each major grid unit in Figure 10.27 corresponds to 1 m, specify the location of point P in polar coordinates.

SOLUTION I begin by reading off the rectangular coordinates: $x = +3.6$ m, $y = +2.4$ m. Then I can use Eqs. 10.1 and 10.2 to convert to polar coordinates:

$$r = +\sqrt{x^2 + y^2} = +\sqrt{(3.6\text{ m})^2 + (2.4\text{ m})^2} = +4.3\text{ m}$$

and

$$\theta = \tan^{-1}\left(\frac{y}{x}\right) = \tan^{-1}\left(\frac{2.4\text{ m}}{3.6\text{ m}}\right) = 34°.$$

So, in polar coordinates point P is at $(r, \theta) = (4.3\text{ m}, 34°)$. ✔

In rectangular coordinates, the position vector of a point P is

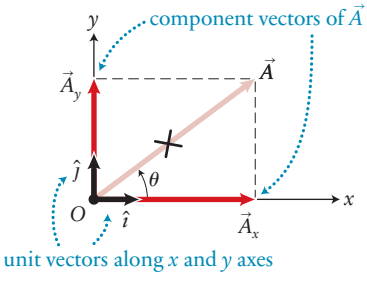
$$\vec{r} \equiv x\,\hat{\imath} + y\,\hat{\jmath}, \tag{10.4}$$

where $\hat{\imath}$ and $\hat{\jmath}$ represent unit vectors along the x and y axes, respectively. The component vectors of $\vec{r}$ are $x\,\hat{\imath}$ and $y\,\hat{\jmath}$. More generally, the decomposition of an arbitrary vector $\vec{A}$ can be written as follows (**Figure 10.28**):

$$\vec{A} = \vec{A}_x + \vec{A}_y = A_x\hat{\imath} + A_y\hat{\jmath}, \tag{10.5}$$

where $\vec{A}_x$ and $\vec{A}_y$ are *component vectors* of $\vec{A}$; A_x and A_y are the x *and* y *components* of $\vec{A}$. Keep in mind that the x and y components, A_x and A_y, are signed numbers—these symbols do *not* represent the magnitudes of the

Figure 10.28 Decomposition of a general vector $\vec{A}$ into component vectors.



QUANTITATIVE TOOLS

component vectors $\vec{A}_x$ and $\vec{A}_y$. Whenever you see a subscript x or y (and later also z) on a symbol that does not have an arrow over it, the quantity represented by that symbol is a component of a vector and carries an algebraic sign.

Using the Pythagorean theorem, we can write the magnitude of $\vec{A}$ as

$$A \equiv |\vec{A}| = +\sqrt{A_x^2 + A_y^2}, \tag{10.6}$$

where the plus sign in front of the square root indicates that only the positive square root applies because magnitude is always a positive quantity. The angle of $\vec{A}$ with respect to the x axis is specified by

$$\tan\theta = \frac{A_y}{A_x}, \tag{10.7}$$

where the signs of A_x and A_y determine in which quadrant θ lies.

The decomposition of vectors is particularly useful when adding or subtracting them. As **Figure 10.29** shows, the x and y components of the vector sum $\vec{R}$ of two vectors $\vec{A}$ and $\vec{B}$ are equal to the sum of the corresponding components of $\vec{A}$ and $\vec{B}$:

$$\vec{R} = (A_x + B_x)\hat{\imath} + (A_y + B_y)\hat{\jmath} \tag{10.8}$$

$$\begin{aligned} R_x &= A_x + B_x \\ R_y &= A_y + B_y. \end{aligned} \tag{10.9}$$

Figure 10.29 Adding vectors by components.

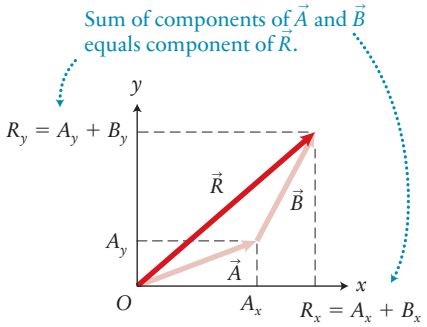
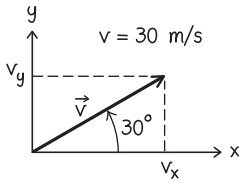


Figure 10.30 Example 10.4.



EXAMPLE 10.4 Speeding ball

A ball is thrown at an angle of 30° above the horizontal at a speed of 30 m/s. Write the ball's velocity in terms of rectangular unit vectors.

❶ **GETTING STARTED** I begin by making a sketch showing the velocity vector $\vec{v}$ and its decomposition in a rectangular coordinate system (**Figure 10.30**). I position the x axis along the horizontal in the direction of motion and the y axis along the vertical. I label the x and y components of the ball's velocity v_x and v_y, respectively.

❷ **DEVISE PLAN** Equation 10.5 tells me that the velocity vector can be written as $\vec{v} = v_x\hat{\imath} + v_y\hat{\jmath}$. To determine the x and y components, I apply trigonometry to the triangle made up by v, v_x, and v_y.

❸ **EXECUTE PLAN** From my sketch I see that $\cos\theta = v_x/v$ and $\sin\theta = v_y/v$, where $v = 30$ m/s and $\theta = 30°$. Therefore $v_x = v\cos\theta$ and $v_y = v\sin\theta$. Substituting the values given for v and θ, I calculate the rectangular components:

$$v_x = (30\text{ m/s})(\cos 30°) = (30\text{ m/s})(0.87) = +26\text{ m/s}$$

$$v_y = (30\text{ m/s})(\sin 30°) = (30\text{ m/s})(0.50) = +15\text{ m/s}.$$

The velocity in terms of unit vectors is thus

$$\vec{v} = v_x\hat{\imath} + v_y\hat{\jmath} = (+26\text{ m/s})\hat{\imath} + (+15\text{ m/s})\hat{\jmath}. \checkmark$$

❹ **EVALUATE RESULT** The x and y components are both positive, as I expect because I chose the direction of the axis in the direction that the ball is moving, and the magnitude of v_x is larger than that of v_y, as it should be for a launch angle that is smaller than 45°. I can quickly check my math by using Eq. 10.6 to calculate the magnitude of the velocity: $v = \sqrt{v_x^2 + v_y^2} = \sqrt{(26\text{ m/s})^2 + (15\text{ m/s})^2} = 30$ m/s, which is the speed at which the ball is launched.

10.8 You navigate a ship from a pier to a buoy 1500 m northeast of the pier. There you sail to a point 300 m south and 700 m east of the buoy. What is your distance from the pier at that point, if you do not move relative to the ship?

10.7 Projectile motion in two dimensions

We can now generalize the description of motion in a plane, starting with the position vector of an object:

$$\vec{r} = x\hat{\imath} + y\hat{\jmath}. \tag{10.10}$$

As in one dimension, the object's instantaneous velocity is

$$\vec{v} \equiv \frac{d\vec{r}}{dt} \equiv \lim_{\Delta t \to 0} \frac{\Delta \vec{r}}{\Delta t}. \tag{10.11}$$

Substituting Eq. 10.10 into Eq. 10.11, we obtain

$$\vec{v} = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}\hat{\imath} + \lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t}\hat{\jmath} = v_x\hat{\imath} + v_y\hat{\jmath}, \tag{10.12}$$

so that

$$v_x = \frac{dx}{dt} \quad \text{and} \quad v_y = \frac{dy}{dt}. \tag{10.13}$$

Likewise, the components of the instantaneous acceleration are

$$a_x = \frac{dv_x}{dt} = \frac{d^2x}{dt^2} \quad \text{and} \quad a_y = \frac{dv_y}{dt} = \frac{d^2y}{dt^2}. \tag{10.14}$$

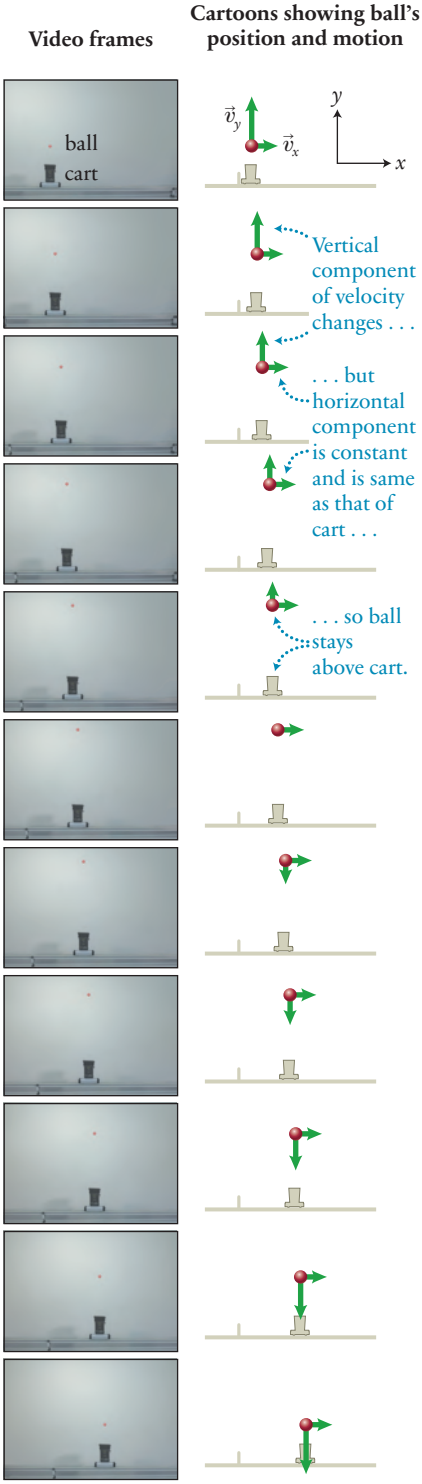
Knowing how to decompose vectors into components allows us to separate motion in a plane into two one-dimensional problems. Consider the motion of a ball launched straight up from a cart moving at constant velocity (**Figure 10.31**). Like the ball dropped from the moving cart in Figure 10.1, the launched ball remains vertically above the cart throughout its trajectory and so falls back into the cart. In the reference frame of the cart, the ball's motion is identical to the one-dimensional projectile motion we studied in Section 3.3. In the Earth reference frame, though, the ball's instantaneous velocity has a component in the horizontal direction in addition to the component in the vertical direction. The resulting two-dimensional projectile motion in the Earth reference frame is the curved trajectory shown in **Figure 10.32**. (For this discussion, ignore air resistance.)

How do we obtain the ball's position at some arbitrary instant $t_f = t_i + \Delta t$? Let us begin by writing the initial condition for the ball's motion in the Earth reference frame at instant t_i. Because the ball's acceleration is always vertically downward, we choose a coordinate system that has a horizontal x axis and a vertical y axis. In the Earth reference frame, the ball's initial velocity is thus

$$\vec{v}_i = v_{x,i}\hat{\imath} + v_{y,i}\hat{\jmath}. \tag{10.15}$$

As viewed from the Earth reference frame, the ball's initial horizontal velocity is equal to the cart's constant velocity. The presence of the horizontal velocity in Eq. 10.15 tells us that, as viewed from the Earth reference

Figure 10.31 Ball launched vertically up from a cart moving at constant velocity in the horizontal direction.



QUANTITATIVE TOOLS

frame, the ball is launched not straight up but at an angle θ relative to the x axis given by

$$\tan\theta = \frac{v_{y,i}}{v_{x,i}}. \tag{10.16}$$

The ball's velocity in the horizontal direction is constant; in the vertical direction, the ball has a constant downward acceleration of magnitude $a_y = -g$. The x component of the ball's instantaneous velocity is not changing. We can obtain the y component at instant t_f from Eq. 3.4:

$$v_{x,f} = v_{x,i} \quad \text{(constant velocity)} \tag{10.17}$$

$$v_{y,f} = v_{y,i} - g\Delta t \quad \text{(constant acceleration)}. \tag{10.18}$$

We can now use the results we derived for one-dimensional projectile motion to determine the position of the ball at any instant. The ball's horizontal position at instant t_f is given by Eq. 2.19, $x_f = x_i + v_{x,i}\Delta t$. For the vertical position we use Eq. 3.8, $x_f = x_i + v_{x,i}\Delta t + \frac{1}{2}a_x(\Delta t)^2$, replacing x by y and then setting $a_y = -g$:

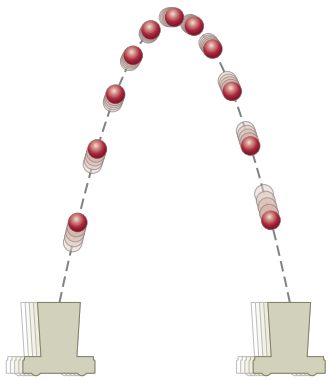
$$x_f = x_i + v_{x,i}\Delta t \quad \text{(constant velocity)} \tag{10.19}$$

$$y_f = y_i + v_{y,i}\Delta t - \tfrac{1}{2}g(\Delta t)^2 \quad \text{(constant acceleration)}. \tag{10.20}$$

These two equations completely determine the trajectory of a projectile launched with initial velocity $\vec{v}_i$ from point (x_i, y_i).

Equations 10.19 and 10.20 give us x_f and y_f as functions of Δt because $x \propto \Delta t$ and $y \propto (\Delta t)^2$. By eliminating Δt from these equations, we obtain a function $y(x)$ that depends quadratically on x and describes a parabola. Hence the trajectory of a projectile whose initial velocity is not vertical is a parabola.

Figure 10.32 Trajectory of the ball in Figure 10.31 as seen from the Earth reference frame.



EXAMPLE 10.5 Position of highest point

The ball of Figure 10.32 is launched from the origin of an xy coordinate system. Write expressions giving, at the top of its trajectory, the ball's rectangular coordinates in terms of the components of its initial speed v_i and the acceleration due to gravity g.

❶ **GETTING STARTED** Because the ball is launched from the origin, $x_i = 0$ and $y_i = 0$. As the ball moves upward, the vertical component of its velocity, v_y, is positive. After crossing its highest position, the ball moves downward, and so now v_y is negative. As the ball passes through its highest position, therefore, v_y reverses sign, so at that position $v_y = 0$.

❷ **DEVISE PLAN** Taking the highest position as my final position and then substituting $v_{y,f} = 0$ into Eq. 10.18, I can determine the time interval Δt_{top} it takes the ball to travel to this position. Once I know Δt_{top}, I can use Eqs. 10.19 and 10.20 to obtain the ball's coordinates at the top.

❸ **EXECUTE PLAN** Substituting $v_{y,f} = 0$ into Eq. 10.18, I get

$$0 = v_{y,i} - g\Delta t_{top}.$$

Solving for Δt_{top} then yields

$$\Delta t_{top} = \frac{v_{y,i}}{g}. \tag{1}$$

Using $x_i = 0$, $y_i = 0$ in Eqs. 10.19 and 10.20, I then calculate the location of the highest position:

$$x_{top} = 0 + v_{x,i}\left(\frac{v_{y,i}}{g}\right) = \frac{v_{x,i}v_{y,i}}{g}$$

and

$$y_{top} = 0 + v_{y,i}\left(\frac{v_{y,i}}{g}\right) - \tfrac{1}{2}g\left(\frac{v_{y,i}}{g}\right)^2$$

$$= \frac{v_{y,i}^2}{g} - \frac{1}{2}\frac{v_{y,i}^2}{g} = \frac{1}{2}\frac{v_{y,i}^2}{g}. \checkmark$$

❹ **EVALUATE RESULT** Because the ball moves at constant velocity in the horizontal direction, the x coordinate of the highest position is the horizontal velocity component $v_{x,i}$ multiplied by the time interval it takes to reach the top (given by Eq. 1). The y coordinate of the highest position does not depend on $v_{x,i}$ as I expect because the vertical and horizontal motions are independent of each other.

QUANTITATIVE TOOLS

EXAMPLE 10.6 Range of a projectile

How far from the launch position is the position at which the ball of Figure 10.32 is once again back in the cart? (This distance is called the *horizontal range* of the projectile.)

❶ **GETTING STARTED** As we saw in Example 3.6, the time interval taken by a projectile to return to its launch position from the top of the trajectory is equal to the time interval Δt_{top} it takes to travel from the launch position to the top. I can therefore use the Δt_{top} expression I found in Example 10.5 to solve this problem.

❷ **DEVISE PLAN** If the time interval it takes the ball to travel from its launch position to the top is Δt_{top}, then the time interval the ball is in the air is $2\Delta t_{top}$. Because the ball travels at constant velocity in the horizontal direction, I can use Eq. 10.19 to obtain the ball's range.

❸ **EXECUTE PLAN** From Example 10.5, I have $\Delta t_{top} = v_{y,i}/g$, and so the time interval spent in the air is $\Delta t_{flight} = 2v_{y,i}/g$. Substituting this value into Eq. 10.19 yields

$$x_f = 0 + v_{x,i}\left(\frac{2v_{y,i}}{g}\right) = \frac{2v_{x,i}v_{y,i}}{g}. \checkmark$$

❹ **EVALUATE RESULT** Because the trajectory is an inverted parabola, the top of the parabola lies midway between the two locations where the parabola intercepts the horizontal axis. So the location at which the parabola returns to the horizontal axis lies a horizontal distance twice as far from the origin as the horizontal distance at the top. In Example 10.5 I found that $x_{top} = v_{x,i}v_{y,i}/g$. The answer I get for the horizontal range is indeed twice this value.



10.9 Suppose a projectile's initial velocity is specified by the initial speed v_i and launch angle θ instead of by its rectangular components as in Eq. 10.15. (*a*) Using Eqs. 10.3, write expressions for the projectile's maximum height and horizontal range in terms of v_i and θ. (*b*) What angle θ gives the greatest value of y_{max}? (*c*) Of x_{max}?

10.8 Collisions and momentum in two dimensions

As we saw in Chapter 5, collisions in one dimension are entirely described by two equations: Eq. 4.17, $\Delta\vec{p} = \vec{0}$, which states that the momentum of the isolated system of colliding objects does not change, and Eq. 5.18, $e \equiv v_{12f}/v_{12i}$, which defines the coefficient of restitution. The same two equations also apply to collisions in two dimensions, but because momentum is a vector, Eq. 4.17 now yields two separate equations: one for the momentum in the x direction and the other for the momentum in the y direction. Consider, for example, a collision in an isolated system of two objects of inertia m_1 and m_2. For this system, Eq. 4.17 yields

$$\Delta p_x = \Delta p_{1x} + \Delta p_{2x} = m_1(v_{1x,f} - v_{1x,i}) + m_2(v_{2x,f} - v_{2x,i}) = 0 \qquad (10.21)$$

$$\Delta p_y = \Delta p_{1y} + \Delta p_{2y} = m_1(v_{1y,f} - v_{1y,i}) + m_2(v_{2y,f} - v_{2y,i}) = 0. \qquad (10.22)$$

If we know the initial velocities of the colliding objects, we have values for $v_{1x,i}$, $v_{1y,i}$, $v_{2x,i}$, and $v_{2y,i}$, but these two equations still contain four unknown quantities: $v_{1x,f}$, $v_{1y,f}$, $v_{2x,f}$, and $v_{2y,f}$. To obtain values for these four unknown quantities, we need four equations. Even if we know the coefficient of restitution of the colliding objects, we have only three equations (Eqs. 10.21, 10.22, and 5.18) and so, unlike collisions in one dimension, we cannot determine the outcome of the collision without some additional information about the final velocities.

QUANTITATIVE TOOLS

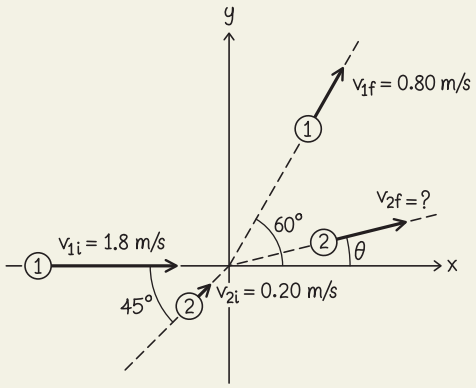
EXAMPLE 10.7 Pucks colliding

Pucks 1 and 2 slide on ice and collide. The inertia of puck 2 is twice that of puck 1. Puck 1 initially moves at 1.8 m/s; puck 2 initially moves at 0.20 m/s in a direction that makes an angle of 45° with the direction of puck 1. After the collision, puck 1 moves at 0.80 m/s in a direction that makes an angle of 60° with its original direction. What are the speed and direction of puck 2 after the collision?

❶ **GETTING STARTED** Because it takes place on a surface, the collision is two-dimensional. I begin by organizing the given information in a sketch (**Figure 10.33**). I let puck 1 move along the $+x$ axis before the collision. Then I add puck 2, which initially moves at an angle of 45° to the x axis. I draw puck 1 after the collision moving upward and to the right along a line that makes an angle of 60° with the $+x$ axis. I don't know

anything about the motion of puck 2 after the collision, and so I arbitrarily draw it moving to the right and upward. I label the pucks and indicate the speeds I know. I need to determine the final speed v_{2f} of puck 2 and the angle θ between its direction of motion after the collision and the x axis.

Figure 10.33



❷ **DEVISE PLAN** Equations 10.21 and 10.22 give the relationship between the initial and final velocities of the two pucks. I know the x and y components of both initial velocities and, after decomposing $\vec{v}_{1f}$, the x and y components of the final velocity of puck 1. Thus I can use Eqs. 10.21 and 10.22 to obtain values for the two components of $\vec{v}_{2f}$. Using these component values, I can calculate v_{2f} and θ.

❸ **EXECUTE PLAN** The two unknown components are $v_{2x,f}$ and $v_{2y,f}$. Solving Eq. 10.21 for $v_{2x,f}$, I get

$$\begin{aligned} v_{2x,f} &= -\frac{m_1}{m_2}(v_{1x,f} - v_{1x,i}) + v_{2x,i} \\ &= -\tfrac{1}{2}[(0.80\ \text{m/s})\cos 60° - (1.8\ \text{m/s})] \\ &\quad + (0.20\ \text{m/s})\cos 45° \\ &= 0.84\ \text{m/s}. \end{aligned}$$

Solving Eq. 10.22 for $v_{2y,f}$, I get

$$\begin{aligned} v_{2y,f} &= -\frac{m_1}{m_2}(v_{1y,f} - v_{1y,i}) + v_{2y,i} = -\frac{m_1}{m_2}v_{1y,f} + v_{2y,i} \\ &= -\tfrac{1}{2}(0.80\ \text{m/s})\sin 60° + (0.20\ \text{m/s})\sin 45° \\ &= -0.21\ \text{m/s}. \end{aligned}$$

Using these values, I obtain the final speed of puck 2 and its direction of motion:

$$v_{2f} = \sqrt{v_{2x,f}^2 + v_{2y,f}^2} = \sqrt{(0.84\ \text{m/s})^2 + (-0.21\ \text{m/s})^2}$$

$$= 0.87\ \text{m/s} \checkmark$$

$$\tan\theta = \frac{v_{2y,f}}{v_{2x,f}} = \frac{-0.21\ \text{m/s}}{0.84\ \text{m/s}} = -0.24$$

or $\theta = -14°$ (or 346°.) ✔

❹ **EVALUATE RESULT** My positive result for $v_{2x,f}$ and negative result for $v_{2y,f}$ tell me that after the collision, puck 2 moves in the positive x direction and the negative y direction. Both of these directions make sense: I can see from my sketch that the x component of the velocity of puck 1 decreases in the collision. The corresponding decrease in the x component of its momentum must be made up by an increase in the x component of the momentum of puck 2. Given that puck 2 initially moves in the positive x direction, it must continue to do so after the collision. In the y direction, I note that because $v_{1y,i}$ is zero, the y component of the momentum of puck 1 increases in the positive y direction by an amount $m_1 v_{1y,f} = m_1[(0.80\ \text{m/s})\sin 60°] = m_1(0.69\ \text{m/s})$. Therefore the y component of the momentum of puck 2 must undergo a change of equal amount in the negative y direction. Because the initial y component of the momentum of puck 2 is smaller than this value, $m_2 v_{2y,i} = 2m_1 v_{2y,i} = 2m_1[(0.20\ \text{m/s})\sin 45°] = m_1(0.28\ \text{m/s})$, puck 2 must reverse its direction of motion in the y direction, yielding a negative value for $v_{2y,f}$, in agreement with what I found.



10.10 Is the collision in Example 10.7 elastic?

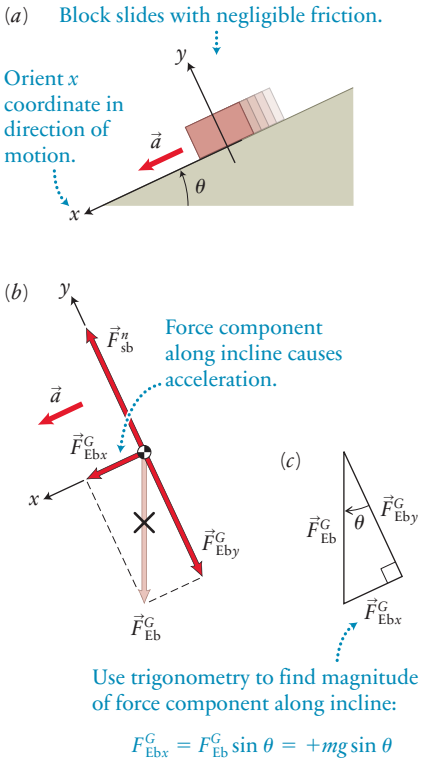
10.9 Work as the product of two vectors

In Section 9.5, I showed that the work done by a constant force on a rigid object is equal to the product of the force and the force displacement:

$$W = F_x \Delta x_F \quad \text{(constant nondissipative force, one dimension)}. \tag{10.23}$$

Work is a scalar, but force and displacement are vectors. What form does the right side of Eq. 10.23 take if the displacement and the force do not point along the same line? Consider the block sliding down an incline

Figure 10.34 Forces exerted on a block sliding down an incline with negligible friction.



shown in **Figure 10.34*a***. The friction between the block and the surface of the incline is negligible. The (nondissipative) force of gravity exerted on the block does not line up with the block's displacement along the incline. In order to calculate the work done on the block, we evaluate the block's change in kinetic energy ΔK while it slides down the incline. Because there are no changes in the internal energy of the block, the work done on it is equal to ΔK.

To calculate ΔK, we must first determine the block's acceleration. Let us choose a coordinate system that has the x axis parallel to the incline and the y axis perpendicular to it, as shown in Figure 10.34*a*. The block is subject to two forces: a normal force of magnitude F^n_{sb} exerted by the incline surface and a downward force of gravity of magnitude F^G_{Eb}, which is shown decomposed into components F^G_{Ebx} and F^G_{Eby} in the free-body diagram in Figure 10.34*b*. (There is no tangential component of the contact force because there is no friction.) The x component of $\vec{F}^G_{Eb}$ is what causes the block to accelerate downward. Using trigonometry, we see from Figure 10.34*c* that

$$F^G_{Ebx} = F^G_{Eb}\sin\theta = +mg\sin\theta, \tag{10.24}$$

where we have substituted mg for the magnitude of the force of gravity exerted on the block (Eq. 8.17) and where the plus sign indicates that this component points in the positive x direction. This equation tells us that the magnitude of the downward acceleration of the block is

$$a_x = \frac{\sum F_x}{m} = \frac{F^G_{Ebx}}{m} = \frac{+mg\sin\theta}{m} = +g\sin\theta. \tag{10.25}$$

This is exactly what we found to be the acceleration of a cart rolling down an incline in Section 3.7 (Eq. 3.20). (We now see that the factor $\sin\theta$ in Eq. 3.20 comes from the decomposition of the force of gravity into components parallel and perpendicular to the incline.)

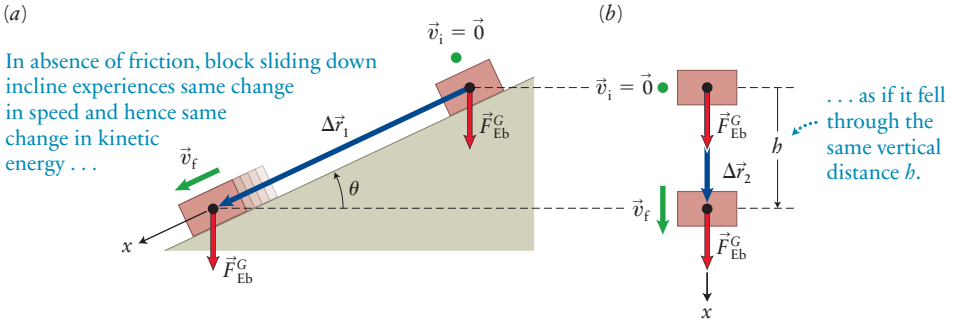
Suppose the block begins at rest on the incline and drops a vertical distance h, as illustrated in **Figure 10.35*a***. If we let the origin coincide with the block's initial position, the block's initial condition is $x_i = 0$, $v_{x,i} = 0$, and its acceleration is given by Eq. 10.25. In order for the block to drop a vertical distance h, its displacement along the incline must be

$$\Delta x = +\frac{h}{\sin\theta}. \tag{10.26}$$

The time the block takes to cover this distance is obtained from Eq. 3.7:

$$\Delta x = v_{x,i}\Delta t + \tfrac{1}{2}a_x(\Delta t)^2 = 0 + \tfrac{1}{2}g\sin\theta(\Delta t)^2, \tag{10.27}$$

Figure 10.35 These blocks gain the same amount of kinetic energy because they descend (without friction) through the same vertical distance h.



and so, from Eqs. 10.26 and 10.27, we obtain

$$(\Delta t)^2 = \frac{2h}{g \sin^2 \theta}. \tag{10.28}$$

The block's final kinetic energy is then

$$\begin{aligned} \tfrac{1}{2}mv_{x,\mathrm{f}}^2 = \tfrac{1}{2}m(v_{x,\mathrm{i}} + a_x\Delta t)^2 &= \tfrac{1}{2}m(0 + g \sin\theta\ \Delta \mathrm{t})^2 \\ &= \tfrac{1}{2}mg^2 \sin^2\theta(\Delta t)^2, \end{aligned} \tag{10.29}$$

where the substitution for $v_{x,\mathrm{f}}$ comes from Eq. 3.4 and that for a_x from Eq. 10.25. Substituting Eq. 10.28 into Eq. 10.29, we obtain

$$\tfrac{1}{2}mv_{x,\mathrm{f}}^2 = \tfrac{1}{2}mg^2 \sin^2\theta\left(\frac{2h}{g \sin^2\theta}\right) = mgh. \tag{10.30}$$

So, because the block's initial kinetic energy is zero, the change in kinetic energy, which is equal to the work done on the block, is

$$W = \Delta K = K_\mathrm{f} - K_\mathrm{i} = \tfrac{1}{2}mv_{x,\mathrm{f}}^2 - 0 = mgh. \tag{10.31}$$

This is the same result we found in Section 7.9 for the change in kinetic energy of a freely falling object! So the work done on the freely falling block in Figure 10.35*b* when it drops a vertical distance h is the same as the work done on the sliding block in Figure 10.35*a* when it drops over the same vertical distance.

Note that we could have obtained the same result for the work done on the block along the incline by calculating the work done on it by *the component of the force of gravity along the incline:*

$$W = F_{\mathrm{Eb}x}^{\mathrm{G}}\Delta x = (+mg \sin\theta)\left(+\frac{h}{\sin\theta}\right) = mgh. \tag{10.32}$$

This means that neither the normal force nor the y component of the force of gravity—both of which are perpendicular to the displacement of the block—does any work on the block. This makes sense because it is the x component of the force of gravity that causes the displacement of the block and that does all the work on it. The forces in the y direction cause no displacement and thus do no work on it.

These results can be stated more compactly by defining a **scalar product** of two vectors. If the angle between two vectors $\vec{A}$ and $\vec{B}$ is ϕ, the scalar product is

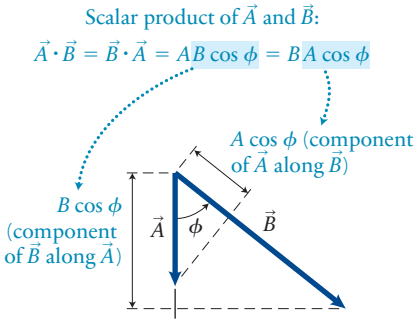
$$\vec{A} \cdot \vec{B} \equiv AB \cos\phi. \tag{10.33}$$

The right side of Eq. 10.33 shows that the scalar product of vectors $\vec{A}$ and $\vec{B}$ is a scalar and is equal to the product of the magnitude of vector $\vec{A}$ ($|\vec{A}| \equiv A$) and the component of $\vec{B}$ along $\vec{A}$ ($B \cos\phi$). Note that the scalar product is commutative, so that the result is independent of the order of multiplication:

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A} = AB \cos\phi = BA \cos\phi, \tag{10.34}$$

which is shown graphically in **Figure 10.36**. Also, as the name suggests, the scalar product of two vectors is a number with units (a scalar), not a vector.

Figure 10.36 The scalar product of two vectors.



Using the definition of scalar product, we can write **work as a scalar product:**

$$W = \vec{F} \cdot \Delta\vec{r}_F \quad \text{(constant nondissipative force)}. \tag{10.35}$$

In one dimension, when the force and the force displacement both point along the same line, Eq. 10.35 reduces to our old expression for work, $W = F_x \Delta x_F$ (Eq. 9.8). Note that if the force and the displacement are in opposite directions, $\phi = 180°$, and so $\cos\phi = -1$ and the work is negative.

Because $\cos 90° = 0$, the work done by a force perpendicular to $\Delta\vec{r}_F$ is zero. So for a block sliding down an incline, the normal force does no work and the work done by the force of gravity on the block is

$$W = \vec{F}^G_{\text{Eb}} \cdot \Delta\vec{r}_F = (mg)\left(\frac{h}{\sin\theta}\right)\cos\phi, \tag{10.36}$$

where θ is the angle of the incline and ϕ the angle between $\vec{F}^G_{\text{Eb}}$ and $\Delta\vec{r}_F$. Because $\phi = \frac{\pi}{2} - \theta$ and $\cos\left(\frac{\pi}{2} - \theta\right) = \sin\theta$, we see as before that the work done by the force of gravity on the block is $W = mgh$.

Figure 10.37 An arbitrary incline can be broken into small straight line segments.

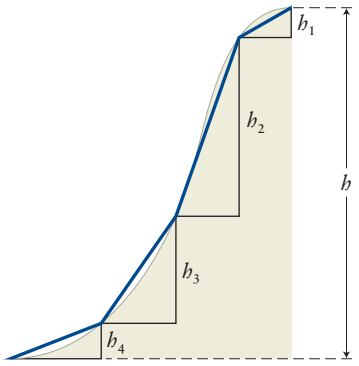


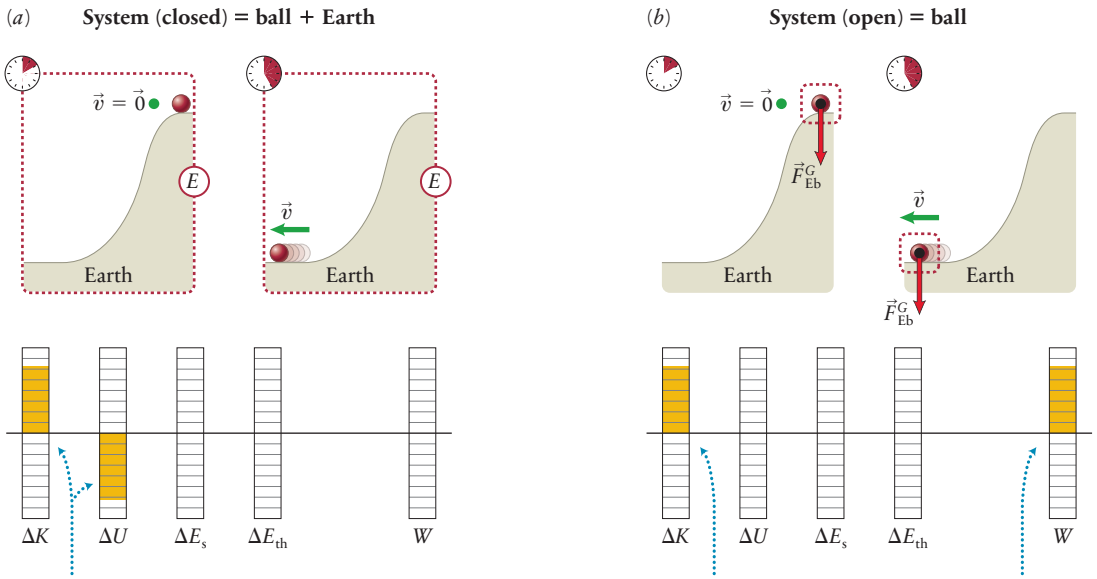
Figure 10.37 shows how we can generalize this result to determine the work done by the gravitational force on an object moving across a surface of any shape. If we approximate the surface by, say, four inclined surfaces, we can apply Eq. 10.36 to each one. The total work done on a block sliding on the surface in Figure 10.37 as it descends from the top to the bottom is

$$W = mgh_1 + mgh_2 + mgh_3 + mgh_4 = mgh. \tag{10.37}$$

To obtain a better approximation we can divide the surface into more, smaller inclined surfaces, but each time we see that the total work done on the block is equal to mgh.

This result also makes sense from an energy perspective. Consider, for example, a ball sliding down the surface of Figure 10.37, and let's assume that there is no friction between the ball and the surface. **Figure 10.38** shows two energy diagrams for this situation. For the closed Earth-ball system

Figure 10.38 Two views of the energy changes that occur when a ball slides down an arbitrary incline with negligible friction.



System gains kinetic energy at expense of its own potential energy; no net change in system's energy.

System gains amount of kinetic energy equal to work done on it by gravitational force.

(Figure 10.38*a*) the change in the ball's kinetic energy ΔK_b and the change in the potential energy of the system add up to zero:

$$\Delta K_b + \Delta U^G = 0. \tag{10.38}$$

From Eq. 7.19, we know that the change in the gravitational potential energy is given by $\Delta U^G = -mgh$, and so

$$\Delta K_b = mgh. \tag{10.39}$$

Next let's consider just the ball by itself as our system (Figure 10.38*b*). There is no change in the ball's internal energy, so the change in the ball's kinetic energy is due to work done on it by the force of gravity: $\Delta K_b = W$. The change in the ball's kinetic energy does not depend on the system chosen, so we see from Eq. 10.39 that the work done on the ball is

$$W = mgh, \tag{10.40}$$

in agreement with Eqs. 10.36 and 10.37. In the process of deriving Eqs. 10.36 and 10.37, however, we derived an expression for the work done by a force that does not point in the same direction as the force displacement (Eq. 10.35).

Let us end by generalizing the result in Eq. 10.35 to a nondissipative force $\vec{F}(\vec{r})$ that is not constant. For such a force we follow the same procedure as in Section 9.7: we divide the force displacement $\Delta\vec{r}_F$ into many small force displacements $\delta\vec{r}_F$, such that the force is approximately constant over each small displacement, and then apply Eq. 10.35 to each small displacement (**Figure 10.39**):

$$W_n \approx \vec{F}(\vec{r}_n) \cdot \delta\vec{r}_{Fn}, \tag{10.41}$$

where $\vec{F}(\vec{r}_n)$ is the force at position $\vec{r}_n$ (at the beginning of each interval, as shown in Figure 10.39). The work done on the system over the entire force displacement $\Delta\vec{r}_F$ is then equal to the sum of the work done over all of the small displacements:

$$W \approx \sum_n W_n = \sum_n \vec{F}(\vec{r}_n) \cdot \delta\vec{r}_{Fn}. \tag{10.42}$$

To get a better approximation, we can divide the force displacement $\Delta\vec{r}_F$ into a larger number of smaller displacements $\delta\vec{r}_{Fn}$, so that the variation of the force over each displacement becomes even smaller. An exact result is obtained if we let $\delta\vec{r}_{Fn}$ approach zero (and the number of small displacements approaches infinity):

$$W = \lim_{\delta\vec{r}_{Fn}\to 0} \sum_n \vec{F}(\vec{r}_n) \cdot \delta\vec{r}_{Fn}, \tag{10.43}$$

which is the definition of the line integral (Appendix B) of the force:

$$W = \int_{\vec{r}_i}^{\vec{r}_f} \vec{F}(\vec{r}) \cdot d\vec{r} \quad \text{(variable nondissipative force)}, \tag{10.44}$$

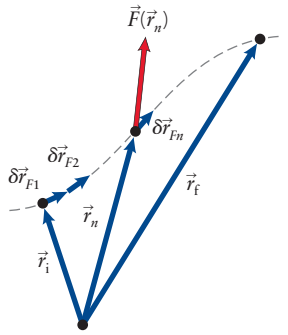
where the line integral should be evaluated over the path traced out by the point of application of the force.

For dissipative forces, we can follow the same procedure of dividing up the displacement to our approach of Section 9.7. Equation 9.27 then replaces Eq. 10.35, yielding for the change in thermal energy caused by the dissipative force:

$$\Delta E_{th} = -\int_{\vec{r}_i}^{\vec{r}_f} \vec{F}(\vec{r}_{cm}) \cdot d\vec{r}_{cm} \quad \text{(variable dissipative force)}. \tag{10.45}$$

Note that, in contrast to Eq. 10.44, this expression contains the position of the center of mass of the system rather than the point of application of the force.

Figure 10.39 To evaluate the work done on a system by a variable force, we divide the force displacement into small pieces.



10.11 When is W in Eq. 10.40 positive, and when is it negative? Is your answer consistent with the definition of positive and negative work given in Section 9.2?

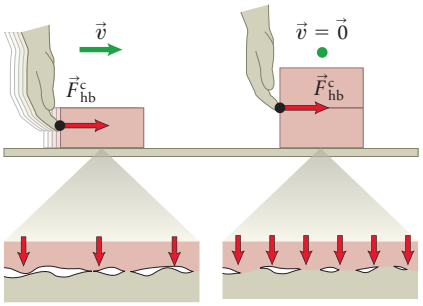
10.10 Coefficients of friction

To describe the motion of objects along surfaces, we need a quantitative expression for the force of friction. Like the normal force, the force of static friction can take on a range of values, depending on the situation.

Suppose you are still struggling with the file cabinet of Checkpoint 10.4. As you increase the strength of your push on the cabinet, the force of static friction opposing your push grows. If you change the direction of your push, the force of friction changes direction as well, always opposing your force. The force of static friction has a maximum value, however, evidenced by the fact that if you push hard enough, the cabinet starts sliding. So the question is: What is the *maximum* force of static friction in a given situation?

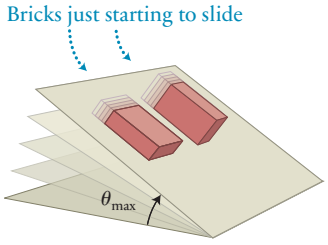
The answer to this question depends on many variables, but we can establish a number of basic facts. For example—this you know from experience—the maximum force of static friction depends on the force with which the surfaces press against each other. An empty file cabinet slides more easily than a full one, for instance, and the double brick in **Figure 10.40** requires twice as large a force to begin sliding as the single brick does. The reason for this dependence on load is shown in Figure 10.40. If you put a brick down on a large flat plate, the *contact area* between the brick and the plate is given by the surface area of the face of the brick that is touching the plate. At the microscopic level, however, surfaces are never completely flat, so the surfaces of the brick and the plate touch only at the tops of protrusions over a combined *effective contact area* that is much smaller than the contact area (typically only about 1/10,000 of it). When the load increases, protrusions on the surfaces flatten, and the effective contact area increases, strengthening the bond between the surfaces.

Figure 10.40 On a rough surface, a single brick slides more easily than a double brick. The harder two surfaces press against each other, the larger the effective contact area, and consequently the stronger the bond between the surfaces.



Measurements show that the angle θ_{max} at which objects begin to slide down an incline (in other words, the angle at which the force of static friction reaches its maximum value) does not depend on the contact area. As illustrated in **Figure 10.41**, two identical bricks begin sliding at the same incline angle regardless of the contact area between the brick and the incline surface. I know it is counterintuitive that the contact area does not determine when the bricks begin to slide, but to understand what is going on, let's go to the microscopic level again. When the brick is oriented so that its largest face is in contact with the inclined surface, the force exerted by the brick on the inclined surface is spread out over a large number of contact points, with the effective contact area having a certain value. When the brick is on its side, there are fewer contact points, but each point now carries a greater share of the load and so flattens out more. This flattening increases the contact area available at each point, with the result that the effective contact area has the same value as when the brick was in the other orientation. Thus the effective contact area doesn't depend on the orientation of the brick—it depends only on the force with which the brick presses against the surface. These observations establish the following important point:

Figure 10.41 Regardless of its orientation, a brick starts sliding at the same angle of incline.



> **The maximum force of static friction exerted by a surface on an object is proportional to the force with which the object presses on the surface and does not depend on the contact area.**

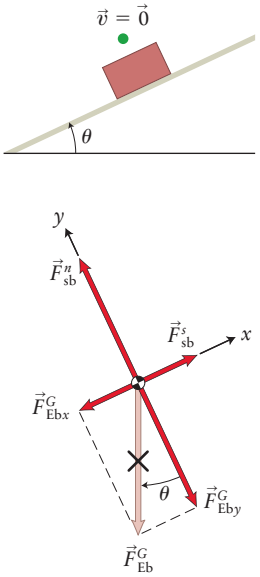
To obtain an expression for the force with which an object presses on an inclined surface, consider **Figure 10.42**. The normal force $\vec{F}^{n}_{sb}$ exerted by

the surface on the block and the normal force $\vec{F}^n_{bs}$ with which the surface of the block presses against the inclined surface (not shown in Figure 10.42) are an interaction pair and therefore equal in magnitude. As we have just determined, the maximum magnitude of the force of static friction is proportional to F^c_{bs} and therefore must also be proportional to F^n_{sb}, $(F^s_{sb})_{max} = \mu_s F^n_{sb}$. For any two surfaces 1 and 2, we can therefore write

$$(F^s_{12})_{max} = \mu_s F^n_{12}. \tag{10.46}$$

The maximum magnitude of the force of static friction between two surfaces is proportional to the magnitude of the normal force exerted by the surfaces on each other. The unitless proportionality constant μ_s is called the **coefficient of static friction.** Its value depends on the two materials in contact and the condition of their surfaces. Values for a few representative materials are given in **Table 10.1**.

Figure 10.42 Free-body diagram for a block at rest on an inclined surface.



10.12 (*a*) A brick slides more easily on ice than on wood. For which combination (brick on ice or brick on wood) is μ_s larger? (*b*) For rubber on dry asphalt, μ_s is about 1. Ignoring any limits due to the runner's skill and condition, what maximum acceleration without slipping can a runner achieve using rubber soles on a dry, horizontal asphalt surface? [Hint: See Figure 10.21.]

The coefficient of static friction expressed in terms of the angle of incline can readily be determined experimentally by raising the incline of Figure 10.42 to the maximum angle at which the block remains at rest. At that angle θ_{max}, the force of static friction $\vec{F}^s_{sb}$ reaches the maximum value given by Eq. 10.46, and because the block remains at rest, the forces along the x axis must add to zero:

$$\Sigma F_x = F^G_{Ebx} + (F^s_{sb})_{max} = 0 \tag{10.47}$$

and so

$$(F^s_{sb})_{max} = -F^G_{Ebx}. \tag{10.48}$$

Table 10.1 Coefficients of friction

Material 1	Material 2	μ_s	μ_k
aluminum	aluminum	1.1–1.4	1.4
aluminum	steel	0.6	0.5
glass	glass	0.9–1.0	0.4
glass	nickel	0.8	0.6
ice	ice	0.1	0.03
rubber	aluminum	0.5	—
rubber	concrete	1.0–4.0	0.8
rubber	rubber	0.84	—
steel	steel	0.8	0.4
steel	brass	0.5	0.4
steel	copper	0.5	0.4
steel	lead	0.95	0.95
wood	wood	0.25–0.5	0.2

The values given are for clean, dry, smooth surfaces.

Because the block has no acceleration in the y direction, we also have

$$\Sigma F_y = F^n_{\text{sb}} + F^G_{\text{Eb}y} = 0, \tag{10.49}$$

and so

$$F^n_{\text{sb}} = -F^G_{\text{Eb}y}. \tag{10.50}$$

Substituting Eqs. 10.48 and 10.50 into 10.46, we get

$$-F^G_{\text{Eb}x} = \mu_s(-F^G_{\text{Eb}y}) \tag{10.51}$$

$$\mu_s = \frac{F^G_{\text{Eb}x}}{F^G_{\text{Eb}y}}. \tag{10.52}$$

Looking at Figure 10.42, we see that $F^G_{\text{Eb}x}/F^G_{\text{Eb}y}$ is the tangent of θ and so

$$\mu_s = \tan\theta_{\text{max}}. \tag{10.53}$$



10.13 (*a*) To what angle θ_{max} does a coefficient of static friction of 1 correspond? (*b*) Put, in turn, a coin, a paper clip, a cell phone, and a comb on the cover of this book and tilt the book to obtain the angle at which each object begins to slide. In each case, how close do you get to the angle calculated in part *a*? What do you conclude from these observations about the coefficients of static friction for these combinations of surfaces?

Equation 10.46 gives only the *maximum* value of the force of static friction. All we know in any situation in which this force is present but has not reached its maximum value is that the magnitude of the force of static friction exerted by surface 1 on surface 2 must obey the condition

$$F^s_{12} \leq \mu_s F^n_{12}, \tag{10.54}$$

where F^n_{12} is the magnitude of the normal force exerted by surface 1 on surface 2. How, then, can you determine the magnitude of the frictional force exerted on an object in a given situation? Inequality 10.54 is not of much help, especially because the normal force can also take on almost any value! The key is to use what you know about the object's motion: If you know its acceleration and inertia, you can use the equation of motion, $\vec{a} = \Sigma\vec{F}/m$, to determine the vector sum of the forces exerted on the object. If you know what other forces are exerted on it, you can extract the values of the normal and frictional forces from this sum. The general procedure for dealing with situations involving frictional forces at surfaces is explained in the Procedure box "Working with frictional forces."

PROCEDURE: Working with frictional forces

1. Draw a free-body diagram for the object of interest. Choose your x axis parallel to the surface and the y axis perpendicular to it, then decompose your forces along these axes. Indicate the acceleration of the object.
2. The equation of motion in the y direction allows you to determine the sum of the y components of the forces in that direction:

   $$\Sigma F_y = ma_y.$$

   Unless the object is accelerating in the normal direction, $a_y = 0$. Substitute the y components of the forces from your free-body diagram. The resulting equation allows you to determine the normal force.
3. The equation of motion in the x direction is

   $$\Sigma F_x = ma_x.$$

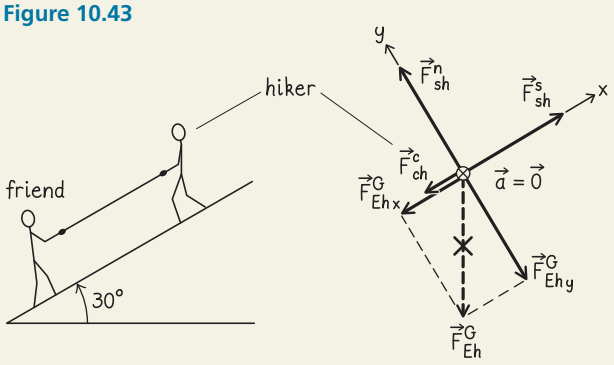
   If the object is not accelerating along the surface, $a_x = 0$. Substitute the x components of the forces from your free-body diagram. The resulting equation allows you to determine the frictional force.
4. If the object is not slipping, the normal force and the force of static friction should obey Inequality 10.54.

EXAMPLE 10.8 Pulling a friend up a hill

A hiker is helping a friend up a hill that makes an angle of 30° with level ground. The hiker, who is farther up the hill, is pulling on a cable attached to his friend. The cable is parallel to the hill so that it also makes an angle of 30° with the horizontal. If the coefficient of static friction between the soles of the hiker's boots and the surface of the hill is 0.80 and his inertia is 65 kg, what is the maximum magnitude of the force he can exert on the cable without slipping?

1 GETTING STARTED I begin by making a sketch and drawing the free-body diagram for the hiker (**Figure 10.43**). I chose an x axis that points up the hill and the y axis perpendicular to it. The forces exerted on the hiker are a force $\vec{F}^{c}_{ch}$ exerted by the cable and directed down the hill (this force forms an interaction pair with the force the hiker exerts on the cable), the force of gravity $\vec{F}^{G}_{Eh}$, and the contact force exerted by the surface of the hill on the hiker. This last force has two components: the normal force $\vec{F}^{n}_{sh}$ and the force of static friction $\vec{F}^{s}_{sh}$. The force of static friction is directed up the hill. If the hiker is not to slip, his acceleration must be zero, and so the forces in the normal (y) and tangential (x) directions must add up to zero.

Figure 10.43



2 DEVISE PLAN As suggested in the Procedure box on page 338, I'll write the equation of motion along the two axes, setting the acceleration in each direction to zero. Then I use Inequality 10.54 to determine the maximum force that the hiker can exert without slipping.

3 EXECUTE PLAN Because the magnitude of the force of gravity exerted on the hiker is mg, I have in the y direction

$$\Sigma F_y = F^{G}_{Eh\,y} + F^{n}_{sh} = -mg\cos\theta + F^{n}_{sh} = ma_y = 0$$

$$F^{n}_{sh} = mg\cos\theta, \qquad (1)$$

where the minus sign in $-mg\cos\theta$ indicates that $F^{G}_{Eh\,y}$ is in the negative y direction. Likewise, I have in the x direction

$$\Sigma F_x = F^{G}_{Eh\,x} + F^{s}_{sh} - F^{c}_{ch}$$

$$= -mg\sin\theta + F^{s}_{sh} - F^{c}_{ch} = ma_x = 0$$

$$F^{s}_{sh} = mg\sin\theta + F^{c}_{ch}. \qquad (2)$$

If the hiker is not to slip, the force of static friction must not exceed its maximum value. Substituting Eqs. 1 and 2 into Inequality 10.54, I get

$$mg\sin\theta + F^{c}_{ch} \le \mu_s(mg\cos\theta)$$

$$F^{c}_{ch} \le mg(\mu_s\cos\theta - \sin\theta).$$

The problem asks me about the magnitude F^{c}_{hc} of the force exerted *by* the hiker *on* the cable, which is the same as the magnitude F^{c}_{ch} of the force exerted *by* the cable *on* the hiker, so

$$F^{c}_{hc} \le mg(\mu_s\cos\theta - \sin\theta).$$

Substituting the values given, I obtain

$$F^{c}_{hc} \le (65\text{ kg})(9.8\text{ m/s}^2)(0.80\cos 30° - \sin 30°)$$

$$= 1.2\times10^2\text{ N}. \checkmark$$

4 EVALUATE RESULT A force of 1.2×10^2 N corresponds to the gravitational force exerted by Earth on an object that has an inertia of only 12 kg, and so the hiker cannot pull very hard before slipping. I know from experience, however, that unless I can lock a foot behind some solid object, it is impossible to pull hard on an inclined surface, and thus the answer I obtained makes sense.

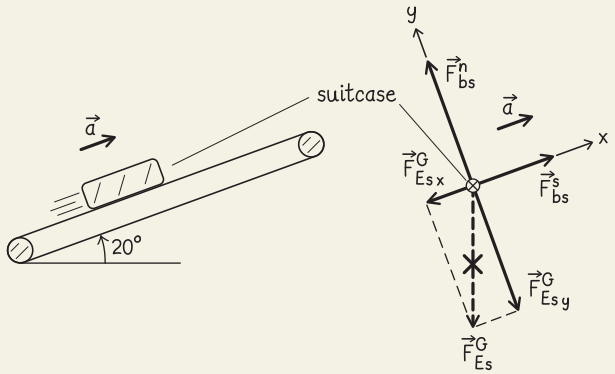
EXAMPLE 10.9 Object accelerating on a conveyor belt

While designing a conveyor belt system for a new airport, you determine that, when a typical suitcase is moving up an incline of 20°, the magnitude of the maximum forward acceleration a rubber belt can give such a suitcase before the suitcase begins slipping is 4.0 m/s^2. What is the coefficient of static friction for a typical suitcase on rubber?

1 GETTING STARTED I begin by making a sketch and drawing a free-body diagram for the suitcase, choosing the x axis along the conveyor belt in the direction of acceleration and the y axis upward and perpendicular to it (**Figure 10.44**). I also draw the upward acceleration of the suitcase along the incline.

2 DEVISE PLAN As in Example 10.8, I write out the equation of motion along both axes. The x component of the

Figure 10.44



(Continued)

acceleration is $a_x = +a$, where the magnitude of the acceleration is $a = 4.0\ \text{m/s}^2$. I can then use Inequality 10.54 to determine the coefficient of static friction.

3 EXECUTE PLAN The equations of motion in the x and y directions are

$$\Sigma F_x = F^G_{\text{Es}x} + F^s_{\text{bs}} = -mg\sin\theta + F^s_{\text{bs}} = ma_x \qquad (1)$$

$$\Sigma F_y = F^G_{\text{Es}y} + F^n_{\text{bs}} = -mg\cos\theta + F^n_{\text{bs}} = ma_y = 0, \qquad (2)$$

with m the inertia of the suitcase. If a_x represents the maximum acceleration at which the suitcase does not slip, the force of static friction must be maximum. Substituting Eq. 10.46 for $(F^s_{\text{bs}})_{\text{max}}$ into Eq. 1, I get

$$-mg\sin\theta + \mu_s F^n_{\text{bs}} = m(+a).$$

Solving this equation for μ_s and substituting $F^n_{\text{bs}} = mg\cos\theta$ from Eq. 2, I obtain

$$\mu_s = \frac{ma + mg\sin\theta}{mg\cos\theta} = \frac{a + g\sin\theta}{g\cos\theta}$$

$$= \frac{(4.0\ \text{m/s}^2) + (9.8\ \text{m/s}^2)\sin 20^\circ}{(9.8\ \text{m/s}^2)(\cos 20^\circ)} = 0.80. \checkmark$$

4 EVALUATE RESULT From Table 10.1, I see that the value I obtained for the coefficient for static friction is close to the coefficient for rubber against rubber and therefore reasonable. I also note that the inertia m of the suitcase, which is not given, drops out of the final result.

Once surfaces slip relative to each other, the force of kinetic friction, like its static counterpart, is proportional to the normal force: It takes a greater force to keep a heavily loaded file cabinet moving than it does to keep an empty one moving. Experiments also show that objects sliding to a stop on a horizontal surface under the influence of kinetic friction undergo a constant acceleration; in other words, the force of kinetic friction between two slipping surfaces is constant and independent of the relative velocity of the two surfaces. In analogy to Eq. 10.46, we can say

$$F^k_{12} = \mu_k F^n_{12}, \qquad (10.55)$$

where μ_k is called the **coefficient of kinetic friction.** The kinetic coefficient μ_k is always smaller than μ_s (or else an object would never slip because it would "stop slipping the instant it starts slipping"). For most surface combinations, the value of μ_k is somewhere between 25% and 100% of the corresponding value of μ_s (Table 10.1).

Note that, unlike the force of static friction, which can take on any value below the maximum limit dictated by μ_s, the force of kinetic friction is constant and therefore always equal to the value given in Eq. 10.55.

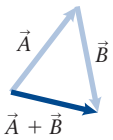
10.14 Imagine putting a single and a double brick on a flat board, as in Figure 10.40, and then slowly raising one end of the board. (*a*) Does the single brick begin to slide before, after, or at the same time as the double brick? (*b*) Once the bricks start sliding, is the acceleration of the double brick larger than, equal to, or smaller than that of the single brick?

QUANTITATIVE TOOLS

Chapter Summary

Vectors in two dimensions (Sections 10.1, 10.2, 10.6, 10.9)

Concepts To add two vectors $\vec{A}$ and $\vec{B}$, place the tail of $\vec{B}$ at the head of $\vec{A}$. Their sum is the vector drawn from the tail of $\vec{A}$ to the head of $\vec{B}$.



To subtract vector $\vec{B}$ from vector $\vec{A}$, reverse the direction of $\vec{B}$ and then add the reversed $\vec{B}$ to $\vec{A}$.

Any vector $\vec{A}$ can be written as

$$\vec{A} = A_x\hat{\imath} + A_y\hat{\jmath}, \tag{10.5}$$

where A_x and A_y are the *components* of $\vec{A}$ along the x and y axes. The x component of $\vec{A}$ is the projection of $\vec{A}$ on the x axis, and the y component is the projection of $\vec{A}$ on the y axis. The components A_x and A_y are signed numbers and can be negative.

Quantitative Tools The magnitude of any vector $\vec{A}$ is

$$A \equiv |\vec{A}| = +\sqrt{A_x^2 + A_y^2} \tag{10.6}$$

and the angle θ that $\vec{A}$ makes with the positive x axis is given by

$$\tan\theta = \frac{A_y}{A_x}. \tag{10.7}$$

If $\vec{R} = \vec{A} + \vec{B}$ is the vector sum of $\vec{A}$ and $\vec{B}$, the components of $\vec{R}$ are

$$\begin{aligned} R_x &= A_x + B_x \\ R_y &= A_y + B_y. \end{aligned} \tag{10.9}$$

The **scalar product** of two vectors $\vec{A}$ and $\vec{B}$ that make an angle ϕ when placed tail to tail is

$$\vec{A}\cdot\vec{B} \equiv AB\cos\phi. \tag{10.33}$$

Projectile motion in two dimensions (Section 10.7)

Concepts For a projectile that moves near Earth's surface and experiences only the force of gravity, the acceleration has magnitude g and is directed downward. The projectile's horizontal acceleration is zero, and the horizontal component of its velocity remains constant.

At the highest point in the trajectory of a projectile, the vertical velocity component v_y is zero, but the vertical component of the acceleration is g and is directed downward.

Quantitative Tools If a projectile is at a position (x, y), its **position vector** $\vec{r}$ is

$$\vec{r} = x\hat{\imath} + y\hat{\jmath}. \tag{10.10}$$

If a projectile has position components x_i and y_i and velocity components $v_{x,\mathrm{i}}$ and $v_{y,\mathrm{i}}$ at one instant, then its acceleration, velocity, and position components a time interval Δt later are

$$a_x = 0$$

$$a_y = -g$$

$$v_{x,\mathrm{f}} = v_{x,\mathrm{i}} \tag{10.17}$$

$$v_{y,\mathrm{f}} = v_{y,\mathrm{i}} - g\Delta t \tag{10.18}$$

$$x_\mathrm{f} = x_\mathrm{i} + v_{x,\mathrm{i}}\Delta t \tag{10.19}$$

$$y_\mathrm{f} = y_\mathrm{i} + v_{y,\mathrm{i}}\Delta t - \tfrac{1}{2}g(\Delta t)^2. \tag{10.20}$$

Collisions and momentum in two dimensions (Section 10.8)

Concepts Momentum is a vector, so in two dimensions changes in momentum must be accounted for by components. This means two equations, one for x and one for y components of momentum. The coefficient of restitution is a scalar and is accounted for by a single equation.

Quantitative Tools

$$\Delta p_x = \Delta p_{1x} + \Delta p_{2x} = m_1(v_{1x,\mathrm{f}} - v_{1x,\mathrm{i}}) + m_2(v_{2x,\mathrm{f}} - v_{2x,\mathrm{i}}) = 0 \tag{10.21}$$

$$\Delta p_y = \Delta p_{1y} + \Delta p_{2y} = m_1(v_{1y,\mathrm{f}} - v_{1y,\mathrm{i}}) + m_2(v_{2y,\mathrm{f}} - v_{2y,\mathrm{i}}) = 0. \tag{10.22}$$

Forces in two dimensions (Sections 10.2, 10.3)

Concepts In two-dimensional motion, the component of the acceleration parallel to the instantaneous velocity changes the speed and the component of the acceleration perpendicular to the instantaneous velocity changes the direction of the instantaneous velocity.

When choosing a coordinate system for a problem dealing with an accelerating object, if possible make one of the axes lie along the direction of the acceleration.

Friction (Sections 10.4, 10.10)

Concepts When two surfaces touch each other, the component of the contact force perpendicular (normal) to the surfaces is called the **normal force** and the component parallel (tangential) to the surfaces is called the **force of friction.**

The direction of the force of friction is such that the force opposes *relative* motion between the two surfaces.

When the surfaces are not moving relative to each other, we have **static friction.** When they are moving relative to each other, we have **kinetic friction.**

The magnitude of the force of kinetic friction is independent of the contact area and independent of the relative speeds of the two surfaces.

Quantitative Tools The maximum magnitude of the force of static friction between any two surfaces 1 and 2 is proportional to the normal force:

$$(F_{12}^s)_{\max} = \mu_s F_{12}^n, \tag{10.46}$$

where μ_s is the unitless **coefficient of static friction.** This upper limit means that the magnitude of the frictional force must obey the condition

$$F_{12}^s \leq \mu_s F_{12}^n. \tag{10.54}$$

The magnitude of the force of kinetic friction is also proportional to the normal force:

$$F_{12}^k = \mu_k F_{12}^n, \tag{10.55}$$

where $\mu_k \leq \mu_s$ is the unitless **coefficient of kinetic friction.**

Work (Sections 10.5, 10.9)

Concepts For a sliding object, the normal force does no work because this force is perpendicular to the direction of the object's displacement.

The force of kinetic friction is a nonelastic force and thus causes energy dissipation.

The force of static friction is an elastic force and so causes no energy dissipation.

The work done by gravity is independent of path.

Quantitative Tools The work done by a constant nondissipative force when the point of application of the force undergoes a displacement $\Delta\vec{r}_F$ is

$$W = \vec{F} \cdot \Delta\vec{r}_F. \tag{10.35}$$

The work done by a variable nondissipative force when the point of application of the force undergoes a displacement is

$$W = \int_{\vec{r}_i}^{\vec{r}_f} \vec{F}(\vec{r}) \cdot d\vec{r}. \tag{10.44}$$

This is the line integral of the force over the path traced out by the point of application of the force.

For a variable dissipative force, the change in the thermal energy is

$$\Delta E_{\text{th}} = -\int_{\vec{r}_i}^{\vec{r}_f} \vec{F}(\vec{r}_{\text{cm}}) \cdot d\vec{r}_{\text{cm}}. \tag{10.45}$$

When an object descends a vertical distance h, no matter what path it follows, the work gravity does on it is

$$W = mgh. \tag{10.40}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

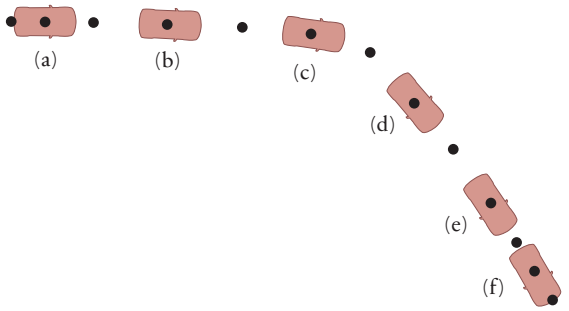
10.1 *Straight* is a relative term

1. Suppose a missile that was launched vertically moves along a trajectory and has a constant vertical speed and a constant horizontal acceleration. What is the shape of this trajectory as seen by a stationary observer? •
2. Riding in a train on horizontal tracks, you notice a boy drop a baseball from the roof of his house nearby. (*a*) What is the shape of the trajectory you see for the ball if the train is moving at constant speed? (*b*) Would you see the trajectory as a straight line if the train were accelerating? ••
3. You are performing an experiment similar to the one shown in Figure 10.1. The ball is dropped from a height of $y_i = 0.200$ m on the pole above the cart just as the pole and cart pass your origin of coordinates. The ball lands inside the cart at the instant the cart passes the location $x = +0.0600$ m. The lab exercise includes, among other things, that you determine the distance traveled by the ball from when it is released until it lands in the cart. Your lab partner immediately says "0.200 m," but the girl at the next table says "0.0600 m." The lab partners just behind you are arguing about whether it is 0.209 m or 0.214 m. You have not done the calculation, but of these four numbers, which could be correct? ••

10.2 Vectors in a plane

4. Is it possible for the acceleration and velocity vectors of an object to *always* be in exactly opposite directions? If no, explain why not. If yes, describe the motion. •
5. A baseball player hits a long fly ball to center field. What can we say about the direction of the ball's acceleration at the highest point of its flight? (Consider air resistance.) •
6. Can an object be accelerated without changing its kinetic energy? •
7. (*a*) Can two vectors of unequal lengths add up to zero? (*b*) What about three vectors of unequal lengths? ••
8. Three swimmers who all swim at the same speed discuss how to cross a river in the shortest amount of time. Swimmer A will swim straight across the river at a right angle to the current. Swimmer B reasons that the current will carry A downstream, meaning that A will cover a greater distance to get across and therefore will take a longer time interval. B says he will aim at an upstream angle such that, allowing for the current, he will reach the other side directly across from where he starts, thus covering the shortest distance and arriving first. Swimmer C, reasoning that the time interval needed for B to cross will be longer than B expects because some of B's effort will be spent battling the current, plans to aim at a downstream angle, so that the current assists rather than opposes him. This way he will be traveling at the highest speed and get across first. Which swimmer gets across first? ••
9. Figure P10.9 shows the position of a regularly flashing light on top of a car (bird's-eye view) as the car moves from (a) to (f). Describe the acceleration of the car at each labeled position based on the information in the figure. ••

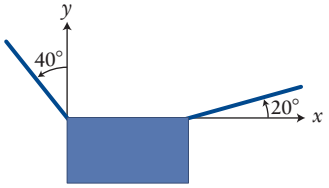
Figure P10.9



10.3 Decomposition of forces

10. You are sweeping the floor with a long-handled broom, with the handle making an acute angle to the horizontal floor. Is it easier to push the broom or pull it? Why? •
11. A sweater hanging on a clothesline makes the line sag in the center. Could you tighten the clothesline enough (assuming whatever material for the line you want) to eliminate the sag completely? ••
12. A sign is suspended by two wires as shown in Figure P10.12. Is the tension in each wire *larger than*, *equal to*, or *smaller than* the gravitational force exerted on the sign by Earth? ••

Figure P10.12



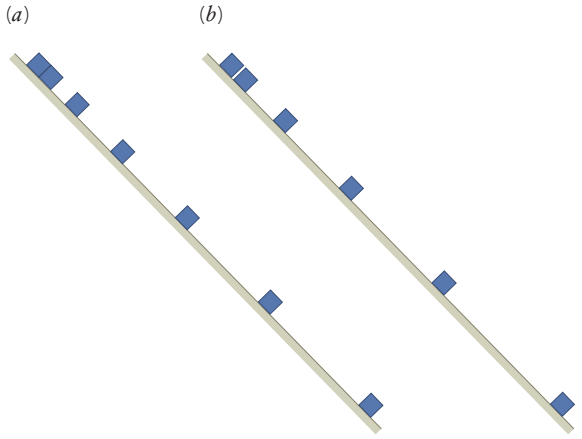
13. A man is crossing a roof that is inclined at an angle of 30° above the horizontal. His path takes him parallel to the roof ridge but 4.0 m (along the roof surface) below the peak. He is supported by two ropes, one attached to an object above each end of the roof. The left rope is attached to a chimney at 1.0 m above the peak of the roof; the right rope is attached to a flagpole at 3.0 m above the peak of the roof. The chimney rope runs parallel to the roof surface, but the flagpole rope does not. Both ropes reach the man at waist level. When the man is halfway across the 10-m-wide roof, describe each force that acts on the man, and identify which have components in each of three perpendicular directions: parallel to the roof ridge, normal to the roof surface, and tangential to the roof surface. •••

10.4 Friction

14. A common parlor trick is to yank a tablecloth hard so that any dishes and glasses on the cloth stay put instead of falling onto the floor. (*a*) Explain how this happens. (*b*) Why, if you pull on the cloth too slowly, do the dishes and glasses get dragged off the table? •
15. A pool ball that is struck gently will roll smoothly across the felt surface of a pool table. However, it is possible to strike the ball in a way that produces "backspin", such that the ball is rotating in the opposite direction to its normal rolling motion as it moves across the table. Describe the nature of the friction force between the table and the ball in these two cases in terms of static or kinetic friction. •

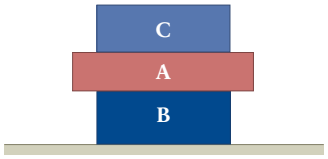
16. Figure P10.16 represents two stroboscopic pictures (taken at the same flash rate) of a block moving along a ramp that exerts a frictional force on the block. In one sketch, the block is released from rest at the top of the ramp and moves downward. In the other sketch, the block is placed at the bottom of the ramp and shoved up the ramp. Which sketch is for the downward motion? Which is for the upward motion? •

Figure P10.16



17. You keep a chalkboard eraser pressed against the chalkboard by using your finger to exert a horizontal force on the back of the eraser. (*a*) Which type of force (call it force A) keeps the eraser from falling? (*b*) What is the magnitude of force A? (*c*) Does the magnitude of force A (which keeps the eraser from falling) depend on how hard you press? (*d*) Now release your finger so that the eraser falls. What happens to the magnitude of force A now? ••
18. In Figure P10.18, all surfaces experience friction. Is the magnitude of the horizontal component of the force exerted by A on B greater than, equal to, or less than the magnitude of the horizontal force exerted by A on C if the system of blocks is (*a*) moving at a constant velocity to the right and (*b*) accelerating to the right due to an external force exerted on B? ••

Figure P10.18



19. A factory worker pulls on a rope attached to a pallet loaded with a crate, dragging the combination at constant speed across a rough factory floor. The tension in the rope acts horizontally on the pallet. Draw a free-body diagram for the pallet and one for the crate when the pallet (*a*) is on level ground and (*b*) is being pulled up an incline. The velocity is constant in both situations. ••

10.5 Work and friction

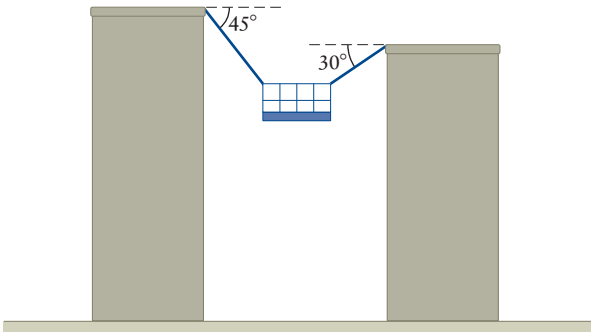
20. What is the direction of the frictional force exerted on a coffee cup (*a*) sitting on a stationary table, (*b*) being pulled to the right across a table, and (*c*) sitting on a table while the table is dragged to the right? •
21. You and a friend are standing on the icy surface of a frozen pond. You are wearing boots with cleats that grip the ice, but your friend has smooth-soled shoes that slip a bit when they move their feet. Discuss how you can both walk across the pond in terms of work and friction. •
22. A resort uses a rope to pull a 65-kg skier up a slope at constant speed. To pull the skier 75 m requires 18 kJ of work to be done by the rope. (*a*) Determine the tension in the rope if the snow is slick enough to allow you to ignore any frictional effects. (*b*) What angle does the slope make to the horizontal? ••
23. A friend claims that her car can accelerate from a stop to 60 mi/h (26.8 m/s) in 5.1 s, but the speedometer is broken. You decide to test her claim by riding with her, and you bring along a small metal washer, a short length of string, a protractor, and a pen that can write on glass. Sitting in the passenger seat, you tie the washer to one end of the string and then let the string-washer combination hang straight down by rolling up the passenger-door window to pinch the free end of the string. While the car is stationary, you draw a line on the window alongside the vertical string. (*a*) How large is the angle the string makes with the line you drew if the car accelerates at the rate your friend claims? (*b*) If an identical washer does not slide on the horizontal dashboard while this is happening, what is the ratio of the frictional force to the gravitational force exerted on this washer? ••
24. You drive a car onto a 100-vehicle ferry, park, and have a short nap. When the ferry arrives at its destination, you start the car and drive off to begin your vacation. Estimate the magnitude of work done by the forces of static friction between the car tires and either the roadway or the ferry surface during each portion of this journey. ••
25. Two children slide down playground slides of identical heights but different angles, one shallow and one steep. (*a*) Suppose the slides are so slippery that the force of friction is negligibly small. Which child, if either, moves faster at the bottom of the slide? Which one, if either, arrives at the bottom first? (*b*) Now take into account the kinetic friction between the child and the slide. Which child, if either, moves faster at the bottom of the slide? ••

10.6 Vector algebra

26. For the vectors $\vec{A} = 2.0\hat{\imath} + 3.0\hat{\jmath}$ and $\vec{B} = -4.0\hat{\imath} + 5.0\hat{\jmath}$, calculate (*a*) $\vec{A} + \vec{B}$ and (*b*) $\vec{A} - \vec{B}$. (*c*) If the direction (but not the magnitude) of $\vec{B}$ is allowed to change by any amount you like, how large can $|\vec{A} + \vec{B}|$ be? How small? •
27. For the vectors $\vec{A} = 4.0\hat{\imath} - 3.0\hat{\jmath}$ and $\vec{B} = -2.0\hat{\imath} + 2.0\hat{\jmath}$, determine (*a*) $\vec{A} - \vec{B}$ and (*b*) $|\vec{A} - \vec{B}|$. •
28. Determine the polar coordinates of vectors (*a*) $\vec{A} = 3.0\hat{\imath} + 2.0\hat{\jmath}$ and (*b*) $\vec{B} = -2.0\hat{\imath} + 2.0\hat{\jmath}$. •
29. Your directions on a scavenger hunt map direct you to walk 55 m north, then 38 m southwest, and finally 12 m east. With the positive x direction being east, what is your displacement (*a*) in Cartesian coordinates and (*b*) in polar coordinates? ••
30. In Figure P10.30, if the tension in the cable attaching the platform to the building on the right is 800 N, what are

(*a*) the tension in the cable attaching the platform to the building on the left and (*b*) the inertia of the platform? ●●

Figure P10.30

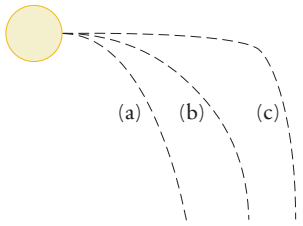


31. You leave your house and walk east for 1.0 h, northeast for 1.5 h, south for 1.0 h, and southwest for 2.5 h, always moving at the same speed. Realizing it is going to get dark soon, you head directly home. How long does it take to walk directly home if your speed stays the same as it was on every leg of the walk? ●●
32. An object's displacement is given by $\vec{r}(t) = (At + Bt^3)\hat{\imath} + (C - Dt^2)\hat{\jmath}$. (*a*) Write the expression for $\vec{v}(t)$ for this motion and the expression for $\vec{a}(t)$. (*b*) Using the values $A = 2.0$ m/s, $B = 0.10$ m/s^3, $C = 5.0$ m, and $D = 0.20$ m/s^2, plot the object's position on a graph that has x as the horizontal axis and y as the vertical axis for instants $t = 0$, 1.0 s, 2.0 s, and 3.0 s. At these positions, draw the velocity and acceleration vectors. ●●
33. Translate the position function $\vec{r}(t) = A\cos(\omega t)\,\hat{\imath} + A\sin(\omega t)\hat{\jmath}$ into polar coordinates, where A and ω are constants. ●●
34. A plane travels in a straight line from position A to position B in 65 min, moving at an average speed of 400 km/h. In a car traveling from A to B, the driver finds that the trip is 600 km long by the route he is forced to take, which goes straight south for a certain distance and then straight west for a longer distance. At what angle south of west does the plane travel? ●●
35. A runner is training on a perfectly circular track with radius 85 m. They run at 7.0 m/s for 1.0 minute. What are the athlete's (*a*) displacement, (*b*) average velocity, and (*c*) magnitude of average velocity? ●●
36. (*a*) In a Cartesian coordinate system that has axes x and y, the coordinates of a certain point P are (x, y). What are the coordinates of P in a Cartesian coordinate system that has axes x' and y' if the origins coincide but the x' axis is tipped above the x axis by an angle φ? (*b*) If the x and y coordinates of P are (5.0 m, 2.0 m) and $\varphi = 30°$, determine (x', y'). ●●●
37. (*a*) Suppose you have two arrows of equal length on a tabletop. If you can move them to point in any direction but they must remain on the tabletop, how many distinct patterns are possible such that the arrows, treated as vectors, sum to zero? [Note: If a pattern cannot be rotated on the tabletop to match another pattern in your list, it is a distinct pattern.] (*b*) Repeat part *a* for three identical arrows, assuming the angle between any adjacent pair must be equal. (*c*) Repeat for four arrows, keeping the angle between neighboring arrows the same throughout the pattern. (*d*) What is the relationship between the number of arrows (vectors) and the number of distinct patterns? ●●●

10.7 Projectile motion in two dimensions

38. An egg is thrown horizontally out of a window 1.5 m off the ground with an initial speed of 8.5 m/s. How far from the building does the egg hit the ground? ●
39. Which is the best representation in Figure P10.39 of the trajectory of a cantaloupe thrown horizontally off a bridge? What is wrong with the other paths? ●

Figure P10.39



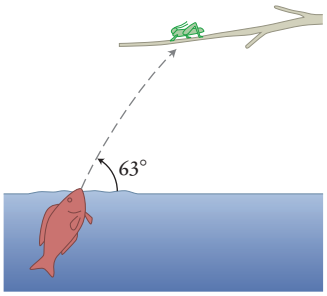
40. A bullet is fired horizontally from a high-powered rifle. At the same instant, a bullet that was resting on top of the rifle falls off. Which bullet hits the ground first? ●
41. A rifle is aimed horizontally at a target 100 m away, and the bullet leaves the rifle barrel at 650 m/s. If the gun is aimed right at the bull's-eye, by how much does the bullet miss the center? ●
42. On a rifle that has a telescopic sight, the telescope is usually not parallel to the barrel of the rifle. The angle the telescope makes with the barrel has to be adjusted for the distance to the target. Explain why. ●●
43. A ball rolls off a 0.65-m high table and lands on the floor. The speed of the ball as it hits the floor is 4.6 m/s. (*a*) How fast was the ball rolling as it left the table? (*b*) What is the horizontal distance from the edge of the table to the ball's landing point? ●●
44. A package is dropped from a helicopter traveling at 15 m/s (horizontally) at an altitude of 200 m, but the parachute attached to the package fails to open. (*a*) How long does it take for the package to hit ground? (*b*) How far does the package travel horizontally before it lands? (*c*) What is the speed of the package just before it lands? ●●
45. The velocity of an object is given in SI units by $\vec{v} = (at - bt^2)\hat{\imath} + c\hat{\jmath}$, with $a = 14$ m/s^2, $b = 10$ m/s^3, and $c = 22$ m/s. (*a*) If the initial position of the object at $t = 0$ is at the origin ($x_i = y_i = 0$), when, if ever, does the object return to the origin? (*b*) When, if ever, is the velocity zero? (*c*) When, if ever, is the acceleration zero? ●●
46. A cannon launches two shells at the same speed, one at 55° above horizontal and one at 35° above horizontal. Which shell, if either, has the longer range? Which shell, if either, is in the air longer? Assume level ground and ignore air resistance. ●●
47. A golf ball is hit from the tee and travels above level ground. Accounting for air resistance, where is the horizontal location of the peak of the flight: at a position that is less than half the range, at a position that is half the range, or at a position that is more than half the range? ●●
48. A cannon fires a shell at an angle such that the shell's initial velocity has a horizontal component of 20 m/s and a vertical component of 35 m/s. Draw the approximate location of the shell at the instants 0, 1.0 s, 2.0 s, 3.0 s, and 4.0 s. At each location, draw the horizontal and vertical components of the velocity. ●●

49. Taking air resistance into account, draw the acceleration vector of a golf ball at five positions along its trajectory from tee to green. Assume for simplicity that the force of air resistance has a constant magnitude. ••

50. How far does a fastball released from the pitcher's hand at a speed of 42 m/s fall as it travels from pitcher to batter, a distance of 18.4 m? (Your answer gives you a good idea of why the pitcher stands on a mound of dirt.) ••

51. The fish in Figure P10.51, peering from just below the water surface, spits a drop of water at the grasshopper and knocks it into the water. The grasshopper's initial position is 0.45 m above the water surface and 0.25 m horizontally away from the fish's mouth. If the launch angle of the drop is 63° with respect to the water surface, how fast is the drop moving when it leaves the fish's mouth? ••

Figure P10.51

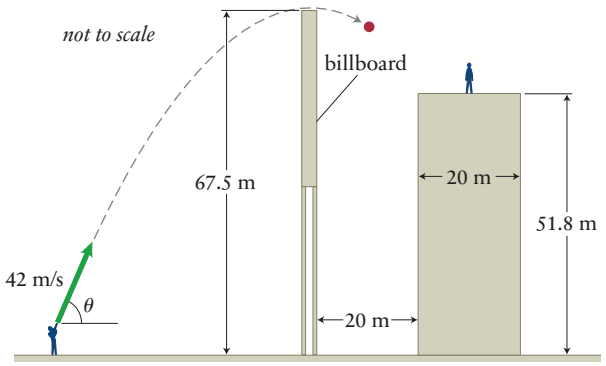


52. A burglar evading the police is running as fast as he can on a building's flat rooftop. At the edge, he sees another flat rooftop across an alley. The horizontal distance across the alley is 8.0 m, the other rooftop is 3.0 m lower than the one on which the burglar is standing, and he can run the 50-m dash in 5.6 s. Can he make the leap when moving at full speed? ••

53. You throw a ball at an angle of 30° above the horizontal at a wall 20 m away. The ball's initial speed is 15 m/s, and it leaves your hand at a height of 1.5 m above the ground. (*a*) How long does the ball take to get to the wall? (*b*) How high up on the wall does it hit? (*c*) What are the horizontal and vertical components of its velocity as it hits the wall? (*d*) When it hits, has it already passed the highest point in its trajectory? ••

54. Figure P10.54 shows a friend standing on the roof of a building that is 51.8 m tall. The roof is square and measures 20 m on a side. You want to shoot a paintball so that it lands on the roof and startles your friend, using a gun that shoots paintballs at a muzzle speed of 42 m/s. The only problem is a slim billboard 67.5 m high between you and the roof, 20 m in front of the building. You position yourself in front of the billboard such that when you hold the gun 1.5 m above the ground and fire, the paintball just barely gets over the billboard at the highest point in its trajectory. (*a*) At what angle θ above the horizontal do you need to shoot the ball to clear the billboard? (*b*) What is your horizontal distance from the billboard? (*c*) How long does the paintball take to move from the highest point in its trajectory to the height of the roof? (*d*) Does the ball strike the roof? (*e*) What is the speed of the ball when it strikes? •••

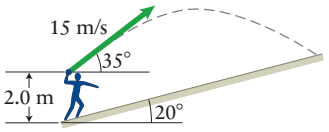
Figure P10.54



55. You throw a textbook to a friend who is at their dormitory window 2.2 m above you. You are standing 4.0 m away from the building and your throw is perfectly made such that when the book arrives at your friend it is travelling exactly horizontally. What is the book's velocity vector as it leaves your hand? ••

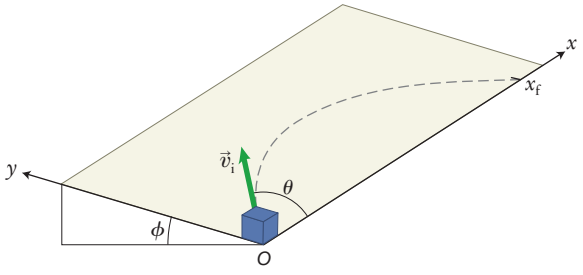
56. You are standing on a slope of 20° to the horizontal and are going to throw a ball at 15 m/s up the incline (Figure P10.56). If you throw the ball at 35° with respect to the horizontal, at what distance up the incline from your feet does it land? Assume the ball leaves your hand directly overhead, 2.0 m above the slope. •••

Figure P10.56



57. A block of inertia m sits at the origin of an xy coordinate system aligned along an incline that makes an angle ϕ with respect to the horizontal (Figure P10.57). You launch the block up the incline, with initial speed v_i directed at an angle θ with respect to the x axis in the plane of the incline. (*a*) What are the magnitude and direction of the block's acceleration expressed in terms of the indicated coordinate system? (*b*) Derive an expression for the block's velocity as a function of time, expressed in unit vector notation for the indicated coordinate system. (*c*) What is the maximum value of the block's displacement Δy along the y axis? (*d*) What is the block's maximum displacement Δx along the x axis? Ignore friction. •••

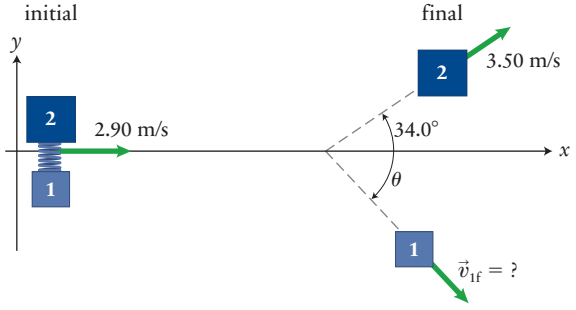
Figure P10.57



10.8 Collisions and momentum in two dimensions

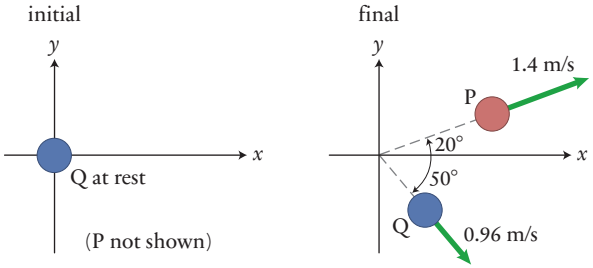
58. A shell explodes prematurely in midair. What happens to the center of mass of the shell? •
59. A cue ball of inertia m is given a speed v before it collides elastically with a full rack of 15 stationary balls in a game of pool. If each of the 16 balls has an inertia m, what are (*a*) the *average* speed of each ball after they fly apart and (*b*) the *average* momentum of each ball after they fly apart? ••
60. Object A, at rest on a low-friction table, is struck a glancing blow by object B. Show that, if the collision is elastic and the two objects have equal inertias, the angle between the final directions of motion of the two objects is 90°. ••
61. Disk 1 (of inertia m) slides with speed 1.0 m/s across a low-friction surface and collides with disk 2 (of inertia $2m$) originally at rest. Disk 1 is observed to bounce off at an angle of 15° to its original line of motion, while disk 2 moves away from the impact at an angle of 55°. (*a*) Calculate the unknown final speed of each disk. (*b*) Is this collision elastic? ••
62. A system of inertia 0.50 kg consists of a spring gun attached to a cart. The system is at rest on a horizontal low-friction track. A 0.050-kg projectile is loaded into the gun, then launched at an angle of 40° with respect to the horizontal plane. With what speed does the cart recoil if the projectile rises 2.0 m at its maximum height? ••
63. A spring ($k = 3800$ N/m) is compressed between two blocks: block 1 of inertia 1.40 kg and block 2 of inertia 2.00 kg. The combination is held together by a string (not shown in Figure P10.63). The combination slides *without spinning* across low-friction ice at 2.90 m/s when suddenly the string breaks, allowing the spring to expand and the blocks to separate. Afterward, the 2.00-kg block is observed to move at a 34.0° angle to its initial line of motion at a speed of 3.50 m/s, while the smaller block moves off at an unknown speed and angle. Neither block is rotating after the separation, and you can ignore the inertias of the spring and the string relative to those of the blocks. (*a*) Determine the velocity of block 1 after the separation. (*b*) Determine the original compression of the spring, $y - y_0$, from its relaxed length y. ••

Figure P10.63



64. Disk P (inertia 0.40 kg) moves at an unknown velocity across a low-friction horizontal surface and collides with disk Q (inertia 0.70 kg), which is initially at rest. After the collision, the two (now slightly dented) disks move apart without spinning. Velocity information is provided in the initial and final top-view diagrams (Figure P10.64). (*a*) What was the initial velocity of disk P? (*b*) What fraction of the initial kinetic energy is converted during the collision? •••

Figure P10.64



10.9 Work as the product of two vectors

65. If the angle between vectors $\vec{A}$ and $\vec{B}$ is 110° and if $A = 1.7$ N and $B = 2.0$ m, what is the value of $\vec{A} \cdot \vec{B}$? •
66. Vector $\vec{A}$ has a magnitude of 3.5 units, and vector $\vec{B}$ has a magnitude of 11 units. If the value of $\vec{A} \cdot \vec{B}$ is 22.4 units squared, what is the angle between $\vec{A}$ and $\vec{B}$? •
67. Vector $\vec{A}$ points in the negative y direction and has a magnitude of 5.0 units. Vector $\vec{B}$ has a positive x component of 3.0 units and a negative y component of 7.0 units. (*a*) What is the angle between the vectors? (*b*) Determine $\vec{A} \cdot \vec{B}$. •
68. A force $\vec{F} = F_x\hat{\imath} + F_y\hat{\jmath}$ with $F_x = 50$ N and $F_y = 12$ N is exerted on a particle as the particle moves along the x axis from $x = 1.0$ m to $x = -5.0$ m. (*a*) Determine the work done by the force on the particle. (*b*) What is the angle between the force and the particle's displacement? ••
69. A force $\vec{F} = F_x\hat{\imath} + F_y\hat{\jmath}$ with $F_x = 3.0$ N and $F_y = 2.0$ N is exerted on a particle that undergoes a displacement $\vec{D} = D_x\hat{\imath} + D_y\hat{\jmath}$ with $D_x = 2.0$ m and $D_y = -2.0$ m. (*a*) Determine the work done by the force on the particle. (*b*) What is the angle between $\vec{F}$ and $\vec{D}$? ••
70. Calculate $\vec{C} \cdot (\vec{B} - \vec{A})$ if $\vec{A} = 3.0\hat{\imath} + 2.0\hat{\jmath}$, $\vec{B} = 1.0\hat{\imath} - 1.0\hat{\jmath}$, and $\vec{C} = 2.0\hat{\imath} + 2.0\hat{\jmath}$. ••
71. You are playing with a toy in which a marble launched by a spring rolls along a track full of loop-the-loops and curves with almost no energy dissipated. You want the marble to have a speed of 0.70 m/s when it arrives at the top of a loop that is 0.30 m higher than the marble's position when you launch it. (*a*) What initial speed must the marble have? (*b*) If the marble has an inertia of 5.0 g and the spring has a spring constant of 13 N/m, how far back do you need to pull the spring on the level launching chute to get the marble to have this initial speed? ••
72. A traveller walks 30 m along a train platform, dragging their suitcase behind them using a strap attached to the top. The traveller's hand is 0.25 m above the top of the suitcase, and the 0.80-m long strap maintains a tension of 120 N as the suitcase is dragged. How much work does the strap do on the suitcase? ••
73. A 20-kg boy slides down a smooth, snow-covered hill on a plastic disk. The hill is at a 10° angle to the horizontal, and the slope is 50 m long. (*a*) If the boy starts from rest, what is his speed at the bottom of the hill? (*b*) His sister then hauls him and the disk from the bottom to the top by means of a rope attached to the disk. What average force must she exert on the rope if the boy is going 0.50 m/s when he reaches the top? ••

74. Two children pull a third child on a tricycle by means of two ropes tied to the handlebars. The combined inertia of the child and tricycle is 35 kg. One child pulls with a constant force of 100 N directed to the right of the straight-ahead direction at an angle of 45°. (*a*) If the second child pulls with a force of 80 N, at what angle to the left of the straight-ahead direction must he pull to make the tricycle move straight ahead? (*b*) With the second child pulling at the angle you calculated in part *a*, what is the acceleration of the tricycle? (*c*) What is the work done on the tricycle by the two ropes if the children pull constantly like this while the tricycle moves 2.0 m? ●●

75. If $\vec{A} = A_x\hat{\imath} + A_y\hat{\jmath}$ and $\vec{B} = B_x\hat{\imath} + B_y\hat{\jmath}$, show that $\vec{A}\cdot\vec{B} = A_xB_x + A_yB_y$. (You may wish to use the fact that the scalar product satisfies the distributive property: $\vec{a}\cdot(\vec{b}+\vec{c}) = \vec{a}\cdot\vec{b} + \vec{a}\cdot\vec{c}$.) ●●●

76. Starting from rest, an intern pushes a 45-kg gurney 40 m down the hall with a constant force of 80 N directed downward at an angle of 35° with respect to the horizontal. (*a*) What is the work done by the intern on the gurney during the 15-m trip? (*b*) How fast is the gurney going when it has moved 15 m? (*c*) How much time elapses during the 15-m journey? Ignore friction. ●●●

77. You know you can provide 500 W of power to move large objects. You need to move a 50-kg safe up to a storage loft, 10 m above the floor. (*a*) With what average speed can you pull the safe straight up? (*b*) How much work do you do on the safe in doing this? (*c*) With what average speed can you pull the safe up a 30° incline? (*d*) How much work do you do on the safe moving if you slide it up the incline to the loft? ●●●

10.10 Coefficients of friction

78. A bartender gives a full mug of beer an initial speed v_{full}, and the mug stops in front of a patron sitting at the end of the bar. This patron later asks for a mug only half full. The bartender stands at the same position as when she sent the full mug to the patron and gives the mug an initial speed v_{half}. If the mug is to again stop in front of the patron, which is true: $v_{\text{half}} < v_{\text{full}}$, $v_{\text{half}} = v_{\text{full}}$, or $v_{\text{half}} > v_{\text{full}}$? ●

79. The heavy crate in Figure P10.79 has plastic skid plates on its bottom surface and a tilted handle attached to one side. Which is easier: pushing the crate or pulling it? Assume your force is exerted along the incline of the handle in either case, and consider friction. ●

Figure P10.79



80. You have lost your skateboard, so you choose to walk around the various sections of the local skate park. The coefficient of static friction between your shoes and the concrete surface is 0.53. What is the angle of the steepest ramp you can stand on without slipping? ●

81. Moving a 51-kg box across a floor, you discover that it takes 200 N of force to get the box moving, and then 100 N keeps it moving at constant speed. What are the coefficients of static and kinetic friction between box and floor? ●

82. Starting from rest, you push your physics book horizontally along a table. Plot the magnitude of the force of friction exerted on the book as a function of the magnitude of the force you exert. Include both static and kinetic cases in the range of force magnitudes you consider. ●●

83. A resort uses a rope to pull a 55-kg skier up a 40° slope at constant speed for 100 m. (*a*) Calculate the tension in the rope if the coefficient of kinetic friction between snow and skis is $\mu_k = 0.20$. (*b*) How much work does the rope do on the skier? (If you did Problem 22, you might be interested in comparing your answers here with those in Problem 22.) ●●

84. When moving on level ground, cross-country skiers slide their skis along the snow surface to stay moving. The coefficients of friction for a given set of skis and given snow conditions can be modified by various types of waxes. In order to move across the snow as fast as possible, (*a*) should you choose a wax that makes the coefficient of static friction between skis and snow as high as possible or as low as possible? (*b*) Should you choose a wax that makes the coefficient of kinetic friction between these two surfaces as high as possible or as low as possible? ●●

85. A janitor is pushing an 11-kg trashcan across a level floor at constant speed. The coefficient of friction between can and floor is 0.10. (*a*) If the janitor is pushing horizontally, what is the magnitude of the force he exerts on the can? (*b*) If he pushes at an angle of 30° down from the horizontal, what must the magnitude of his pushing force be to keep the can moving at constant speed? ●●

86. A hockey puck on the ice starts out moving at 10.50 m/s but after 40.00 m has slowed to 10.39 m/s. (*a*) What is the coefficient of kinetic friction between ice and puck? (*b*) On this same ice, what is the speed after 40.00 m when an identical puck is glued on top of the first one and the combination again starts at 10.50 m/s? ●●

87. The coefficient of kinetic friction between tires and dry pavement is about 0.80. Assume that while traveling at 27 m/s you lock your brakes and as a result the only horizontal force on the car is the frictional one. (*a*) How many seconds does it take you to bring your car to a stop? (*b*) If the road is wet and the coefficient of kinetic friction between tires and pavement is only 0.25, how long does it take? (*c*) How far do you travel in this time interval? ●●

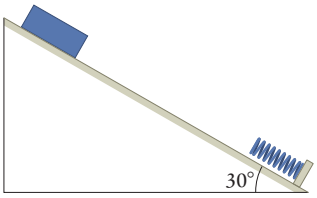
88. A block of inertia m is placed on an inclined plane that makes an angle θ with the horizontal. The block is given a shove directly up the plane so that it has initial speed v, and the coefficient of kinetic friction between the block and the plane surface is μ_k. (*a*) How far up the plane does the block travel before it stops? (*b*) If the coefficient of static friction between block and surface is μ_s, what maximum value of θ allows the block to come to a halt somewhere on the plane and not slide back down? ●●

89. A 1.0-kg block on a horizontal tabletop is pushed against the free end of a spring (the other end is attached to a wall) until the spring is compressed 0.20 m from its relaxed length. The spring constant is $k = 100$ N/m, and the coefficient of kinetic friction between block and tabletop is 0.20. When the block is released from being held against the compressed spring, how far does the block travel before coming to rest? ●●

90. A man exerts a constant force to pull a 50-kg box across a floor at constant speed. He exerts this force by attaching a rope to the box and pulling so that the rope makes a constant angle of 36.9° above the horizontal. The coefficient of kinetic friction for the box-floor interface is $\mu_k = 0.10$. What is the work done by the man as he moves the box 10 m? ••

91. An inclined plane that makes an angle of 30° with the horizontal has a spring of spring constant 4500 N/m at the bottom (Figure P10.91). A 2.2-kg block released near the top of the plane moves down the plane and compresses the spring a maximum of 0.0240 m from its relaxed length. (*a*) Ignoring any frictional effects, calculate the distance the leading edge of the block travels from the instant the block is released to the instant it briefly stops against the compressed spring. (*b*) Now suppose there is friction between the block and the surface of the plane, with $\mu_k = 0.10$. If the block is again placed somewhere on the plane and allowed to slide down and compress the spring by the same 0.0240 m, how far does the leading edge of the block travel now? ••

Figure P10.91



92. You push down on a book of inertia m resting on a table with a force directed at an angle θ away from vertical. The coefficient of static friction between book and table is μ_s. If θ is not larger than some critical value, you cannot get the book to slide no matter how hard you push. What is that critical value? •••

93. You push a 2.0-kg block up a ramp by exerting a 100-N force directed parallel to the ramp, which is at a 30° angle to the horizontal. (*a*) Ignoring any effects due to friction, calculate the block's speed when you have pushed it 2.0 m if its upward speed when you began pushing was 2.0 m/s. (*b*) Now consider friction. If the coefficient of kinetic friction between block and ramp is 0.25, what is the block's speed after you have pushed it 2.0 m? (*c*) With this value of μ_k, how much farther up the ramp does the block move once you stop pushing after the 2.0 m? •••

94. A platform that rolls on wheels and has inertia m_p is attached to a wall by a horizontal spring of spring constant k. A load of inertia m_ℓ sits on the platform, and the coefficient of static friction between platform and load is μ_s. If you pull the platform away from the wall so that the spring is stretched a distance x from its relaxed length and then let go, the platform-spring combination bounces back and forth. (*a*) At what position(s) in the motion is the load most likely to fall off? (*b*) If you want the load to stay in place held only by friction, what is the maximum distance x_{max} you can stretch the spring from its relaxed length and have the load stay on when you release the platform-spring combination? •••

95. You have to specify the power output of a motor for a ski tow rope that will carry 20 skiers at a time. The grade of the ski slope is 32° above horizontal, and the average coefficient of kinetic friction between skis and snow is 0.12. If the tow speed is to be 3.0 m/s and each skier has an inertia of 60 kg (a reasonable average value), what must the minimum power output of the motor be? •••

96. A spring of spring constant k is attached to a support at the bottom of a ramp that makes an angle θ with the horizontal. A block of inertia m is pressed against the free end of the spring until the spring is compressed a distance d from its relaxed length. Call this position A. The block is then released and moves up the ramp until coming to rest at position B. The surface is rough from position A for a distance $2d$ up the ramp, and over this distance the coefficient of kinetic friction for the two surfaces is μ. Friction is negligible elsewhere. What is the distance from A to B? •••

Additional Problems

97. A sports car skids to a stop, leaving skid marks 290 m long. If the coefficient of kinetic friction between tires and pavement is 0.50, how fast was the car going before the skid? •

98. If you were to throw a penny horizontally as hard as you could from the top floor of the Empire State Building, about how far from the base would it land? •

99. Plot the range of a projectile as a function of the launch angle above horizontal. •

100. Watering the garden, you need to have the water reach farther and instinctively raise the hose nozzle to increase the range of the water stream. You realize that increasing the angle the nozzle makes with the ground increases the time interval during which the water is in flight, thereby letting the horizontal motion continue longer. This maneuver works only up to a point, however. Why? ••

101. A particle initially at the origin of an xy coordinate system and having an initial velocity of $(40\ \text{m/s})\hat{\imath}$ experiences a constant acceleration of $\vec{a} = a_x\hat{\imath} + a_y\hat{\jmath}$, with $a_x = -1.0\ \text{m/s}^2$ and $a_y = -0.50\ \text{m/s}^2$. (*a*) What is the maximum x coordinate of the particle's motion? (*b*) What is its velocity when it reaches this position? (*c*) What is its y coordinate at this instant? ••

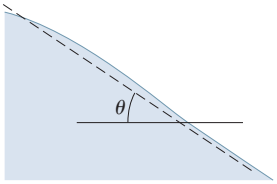
102. Describe the properties of two vectors for which (*a*) $\vec{A} + \vec{B} = \vec{A} - \vec{B}$ and (*b*) $|\vec{A}| + |\vec{B}| = |\vec{A} + \vec{B}|$. ••

103. Three forces are exerted on a 2.00-kg block initially at rest on a slippery surface: a 100-N force directed along the positive x axis, a 50.0-N force that makes an angle of 30.0° with the positive x axis, and a 144-N force that makes an angle of 190° with the positive x axis. (Both angles are measured as viewed from above.) (*a*) What is the vector sum of the forces exerted on the block? (*b*) What is the work done on the block in 10.0 s? ••

104. An airline pilot begins a trip to Duluth from an airport located 1500 km south of Duluth. Her air speed is 260 m/s, but a wind blows from west to east at 40 m/s that takes her off course if she flies directly north. She has the choice of heading slightly westward so that the wind causes her to end up at Duluth or heading due north until she reaches a position directly east of Duluth and then heading west into the wind to Duluth. Which route takes a shorter time interval? ••

105. The coefficient of static friction between layers of snow is 3.7 under ideal conditions. What is the steepest angle a snow field can form (θ in Figure P10.105) before there is an avalanche? ●●

Figure P10.105

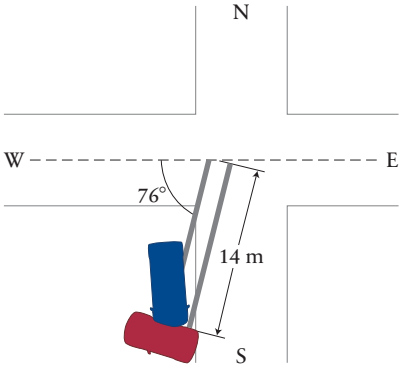


106. Having seen in Example 10.6 that there is a simple expression for the horizontal range of a projectile across level ground, you begin to wonder whether there is an expression for the horizontal range of an object that is thrown with initial speed v_i at an angle θ above the horizontal from the edge of a cliff of height h. ●●● CR

107. You have inherited property in Vermont that would make an excellent ski resort. One of the slopes has a cliff on the other side of the hill, and this gives you a money-saving idea. Instead of a chair lift or motorized tow rope, you decide to attach a pulley to the top of the cliff and then drape the tow rope over the pulley, with one end of the rope temporarily secured at the base of the ski slope and a counterweight attached to the end that hangs over the pulley at the cliff edge. The plan is to release the rope and pull two skiers (with the inertia for the pair kept between 100 kg and 200 kg) up the 400-m slope, which has an incline angle of 35° and a coefficient of kinetic friction between skis and snow no larger than 0.10. You guess that customers get nervous if they move faster than 5.0 m/s, and you wonder if a single counterweight can satisfy these constraints. ●●● CR

108. You are investigating a rural accident. The evidence at the scene suggests that a car was traveling west when struck by a van traveling south. The crushed vehicles stuck together, then came to rest at the end of skid marks of length 14 m along a 76° line as shown in Figure P10.108. (The figure is a top view.) The ground is flat with no curbs, and you estimate a coefficient of kinetic friction $\mu_k = 0.70$ for the crushed vehicles along the path of the skid. You look up the inertias of these model vehicles and find that $m_c = 1400$ kg and $m_v = 6500$ kg. You know that the speed limit is 35 mi/h on both roads, which allows you to decide about driver citations. ●●● CR

Figure P10.108



109. A 0.45-kg soccer ball is kicked at an angle θ above the ground with initial speed v_i. (*a*) Determine the height of the ball as a function of the horizontal distance it travels down the field. (*b*) For $\theta = 30°$, what initial speed must the ball have in order to land 48 m from where it is kicked? (*c*) If the foot is in contact with the ball for 0.15 s, what is the average force exerted by the foot on the ball? ●●●

110. A ballistic pendulum is a device for measuring bullet speeds. One of the simplest versions consists of a block of wood hanging from two long cords. (Two cords are used so that the bottom face of the block remains parallel to the floor as the block swings upward.) A 9.5-g bullet is fired into a ballistic pendulum in which the block has an inertia of 5.0 kg, and the block rises 60 mm above its initial position. (*a*) What is the speed of the bullet just before it hits the block? (*b*) How much energy is dissipated in the collision? What forms does this dissipated energy take? ●●●

111. A 0.010-kg bullet traveling at 300 m/s hits a 2.0-kg ballistic pendulum like the one described in Problem 110. However, the block is not thick enough for this bullet, and the bullet passes through the block, exiting with one-third of its original speed. How high above its original position does the block rise? ●●●



11

Motion in a Circle

11.1 Circular motion at constant speed
11.2 Forces and circular motion
11.3 Rotational inertia

11.4 Rotational kinematics
11.5 Angular momentum
11.6 Rotational inertia of extended objects

CONCEPTS

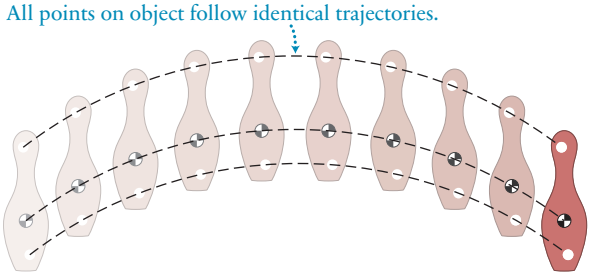
QUANTITATIVE TOOLS

The motion we have been dealing with so far in this text is called **translational motion** (Figure 11.1*a*). This type of motion involves no change in an object's orientation; all the particles in the object move along identical parallel trajectories. During **rotational motion,** which we begin to study in this chapter, the orientation of the object changes, and the particles in an object follow different circular paths centered on a straight line called the *axis of rotation* (Figure 11.1*b*). Generally, the motion of rigid objects is a combination of these two types of motion (Figure 11.1*c*). (Recall that when forces act on a rigid object, we can ignore deformation.) As we shall see in Chapter 12 this combined motion can be broken down into translational and rotational parts that can be analyzed separately. Because we already know how to describe translational motion, knowing how to describe rotational motion will complete our description of the motion of rigid objects.

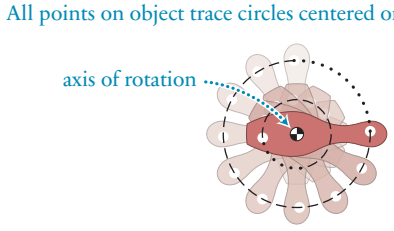
As Figure 11.1*b* shows, each particle in a rotating object traces out a circular path, moving in what we call *circular motion*. We therefore begin our analysis of rotational motion by describing circular motion. Circular motion occurs all around us. A speck of dust stuck to a spinning top, a stone being whirled around on a string, a person on a Ferris wheel—all travel along the perimeter of a circle, repeating their motion over and over. Circular motion takes place in a plane, and so in principle we have already developed all the tools required to describe it. To describe circular and rotational motion we shall follow an approach that is analogous to the one we followed for the description of translational motion. Exploiting this analogy, we can then use the same results and insights gained in earlier chapters to introduce a third conservation law.

Figure 11.1 Translational and rotational motion of a rigid object.

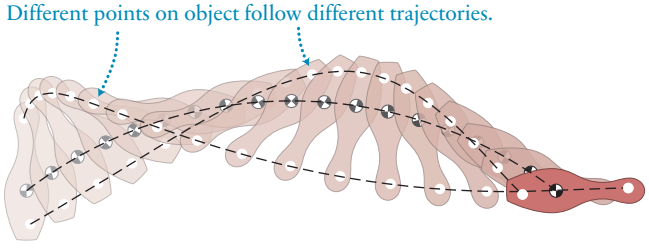
(*a*) Translational motion



(*b*) Rotational motion



(*c*) Combined translation and rotation



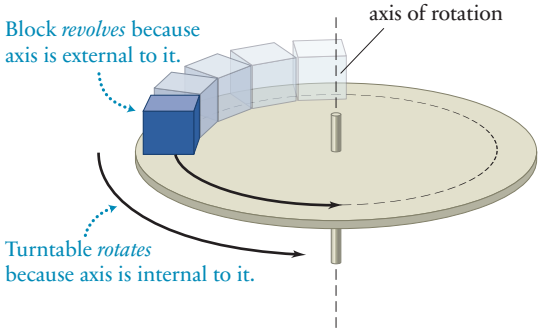
11.1 Circular motion at constant speed

Figure 11.2 shows two examples of circular motion: a block dragged along a circle by a rotating turntable and a puck constrained by a string to move in a circle. The block and puck are said to *revolve* around the vertical axis through the center of each circular path. Note that the axis about which they revolve is external to the block and puck and perpendicular to the plane of rotation. This is the definition of *revolve*—to move in circular motion around an *external* center. Objects that turn about an *internal* axis, such as the turntable in Figure 11.2*a*, are said to *rotate*. These two types of motion are closely related because a rotating object can be considered as a system of an enormous number of particles, each revolving around the axis of rotation.

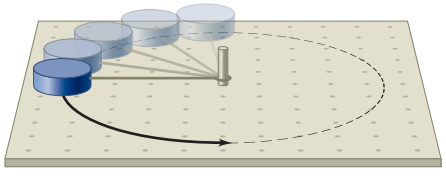
Figure 11.3 shows an overhead view of a puck moving counterclockwise along the arc of a circle. The puck's average velocity points in the same direction as the puck's displacement $\Delta\vec{r}$, which points from the puck's initial position to its final position. As we consider the

Figure 11.2 Examples of circular motion.

(*a*) Block revolves on rotating turntable

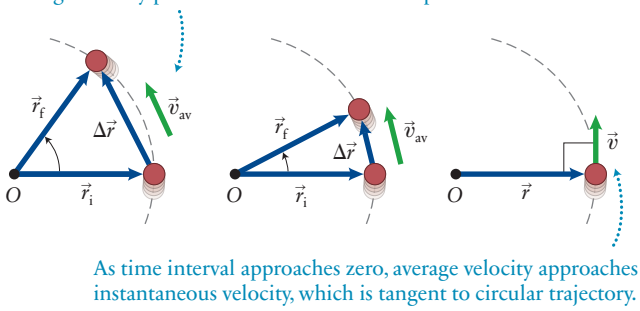


(*b*) Tethered puck revolves on air table



CONCEPTS

Figure 11.3 Average and instantaneous velocity for an object in circular motion.



motion for increasingly smaller time intervals, the angle between the position vectors $\vec{r}_i$ and $\vec{r}_f$ measured from the center of the circular trajectory approaches zero and $\Delta\vec{r}$ becomes tangent to the circular trajectory. This is true all along the trajectory, and so we can say:

> **The instantaneous velocity $\vec{v}$ of an object in circular motion is always perpendicular to the object's position $\vec{r}$ measured from the center of the circular trajectory.**

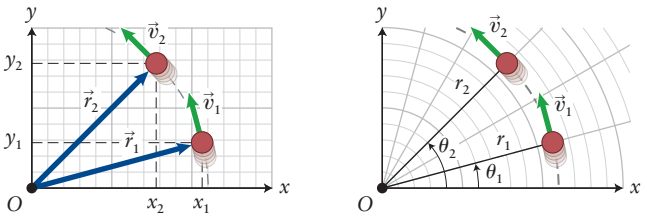
In the first part of this chapter, we study only objects in **circular motion at constant speed;** that is to say, the magnitude of their instantaneous velocity doesn't change.

Because the motion lies entirely in a plane, we can use Cartesian coordinates to describe the puck's position and velocity at any instant. In **Figure 11.4a**, for example, the puck's position is specified by (x_1, y_1) at instant t_1 and by (x_2, y_2) at instant t_2. However, because the magnitude of the position vector $\vec{r}$ is always equal to the radius of the circle, we can also give the magnitude r and the direction of this vector to specify the puck's position, as indicated by the angles θ_1 and θ_2 in Figure 11.4*b*. To specify the direction, we define the puck's **rotational coordinate,** which is a unitless number that increases by 2π for each circle completed by the revolving puck.* The rotational coordinate is denoted by the Greek letter ϑ (script theta); for example, $\vartheta = \pi$ refers to the position of the puck half a revolution after starting from the origin. The rotational coordinate is related to *polar angle* θ as shown in **Figure 11.5a**. Unlike the rotational coordinate, however, the polar angle is not a unitless number but rather expressed in units of revolutions, degrees, or radians.

To measure ϑ we need to choose a direction of increasing ϑ and a zero, just as we need to specify a direction of increasing x and an origin to measure position along an axis. We shall indicate our choice of rotational coordinate system, which will generally be dictated by the motion under consideration, by drawing a curved arrow around the axis of rotation, as shown in Figure 11.5*b*.

Figure 11.4 The position of an object in circular motion can be given in Cartesian or polar coordinates.

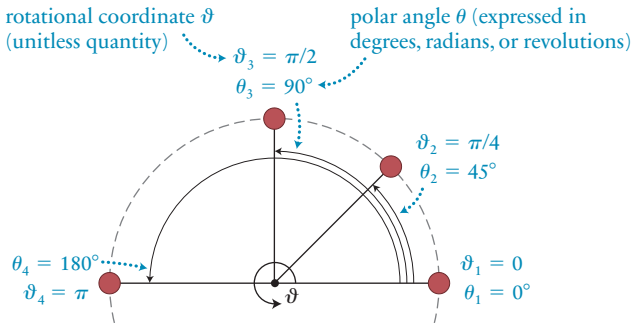
(*a*) Cartesian (x, y) coordinates (*b*) Polar (r, θ) coordinates



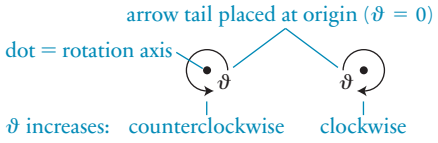
*The adjective *rotational* is used for both rotation and revolution.

Figure 11.5 The rotational coordinate specifies the location of an object on a circle.

(*a*) Relationship between rotational coordinate and polar angle



(*b*) Symbol used in this text to specify rotational coordinate system



EXERCISE 11.1 Rotational coordinates

Determine the rotational coordinate of the following objects for the indicated choices of rotational coordinate system. (*a*) The tip of the hour hand of a clock at 4:30 p.m., when zero is at noon on the same day and ϑ increases in the clockwise direction. (*b*) A faucet after being turned 3/4 of a turn in a clockwise direction, starting from zero and with ϑ increasing in the counterclockwise direction. (*c*) The object in Figure 11.4 if $\theta = 78.3°$, when zero is at $\theta = 0°$ and ϑ increases in the counterclockwise direction.

SOLUTION (*a*) The angle between noon and 4:30 p.m. is 135°, which is $(135°)/(360°) = 0.375$ of one revolution. Because the rotation takes place in the positive direction and because ϑ increases by 2π for each revolution, $\vartheta = (2\pi)(+0.375) = +2.36$. Note that there are no units because ϑ is a unitless quantity. ✔

(*b*) $\vartheta = (2\pi)(-3/4) = -\frac{3}{2}\pi$. ✔

(*c*) $\vartheta = (2\pi)(+78.3°)/(360°) = +1.37$. ✔

CONCEPTS

We shall represent the rate at which an object's rotational coordinate changes by ω_ϑ (Greek omega). The subscript ϑ indicates that the sign of this quantity depends on our choice of direction for increasing ϑ, just as the sign of v_x (the rate at which an object's x coordinate changes) depends on our choice of x axis. Indeed, just as v_x is the component of a velocity vector $\vec{v}$, ω_ϑ is a component of a *rotational velocity vector* $\vec{\omega}$, which we shall discuss in Chapter 12. For convenience we shall refer to ω_ϑ as the **rotational velocity,** although strictly speaking it is the ϑ component of $\vec{\omega}$.

The magnitude of the rotational velocity is the **rotational speed,** which we denote by ω just as v denotes speed: $\omega = |\vec{\omega}| = |\omega_\vartheta|$. The time interval it takes an object in circular motion at constant speed to complete one revolution is called the **period** and is denoted by the letter T.

For example, if the puck in Figure 11.4 completes three revolutions in the negative ϑ direction in 2.0 s, its average rotational velocity for those three revolutions $(\Delta\vartheta = -3(2\pi) = -6\pi)$ is $\omega_{\vartheta,\text{av}} = \Delta\vartheta/\Delta t = (-6\pi)/(2.0\text{ s}) = -9.4\text{ s}^{-1}$; its average rotational speed is 9.4 s^{-1}; and its period is $T = (2.0\text{ s})/3 = 0.67$ s. Note that the units of rotational velocity and rotational speed are s^{-1} (*inverse second*).

The rotational coordinate of an object in circular motion at constant speed increases linearly with time. If it takes a puck in circular motion at constant speed 1.0 s to go from $\vartheta = 0$ to $\vartheta = +0.80$, then it also takes the puck 1.0 s to go from $\vartheta = +0.80$ to $\vartheta = +1.6$, and the object's rotational velocity is constant at $\omega_\vartheta = (+1.6 - 0.80)/(1.0\text{ s}) = +0.80\text{ s}^{-1}$.

What is the relationship between an object's speed and its rotational velocity? **Figure 11.6** shows two pucks in circular motion. The rotational coordinate of both pucks changes by the same amount $\Delta\vartheta$ during a time interval Δt, and so both have the same rotational velocity. Their speeds, however, are not the same because the puck that moves along the larger circle travels a greater distance along its trajectory than does the puck that moves along the smaller circle.

Figure 11.6 Two pucks in circular motion at constant speed can have different speeds even though their rotational velocity (the rate at which ϑ changes) is the same.

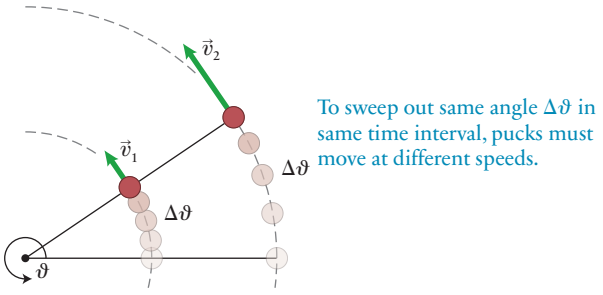
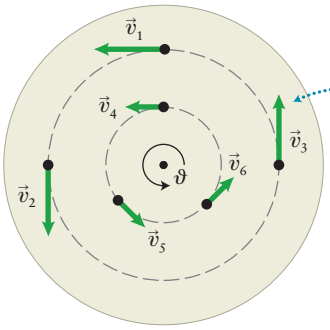


Figure 11.7 Instantaneous and rotational velocities for points on a rotating disk.



The advantage of using rotational velocity to describe rotational motion over using the ordinary velocity $\vec{v}$ becomes clear when we look at the rotating disk in **Figure 11.7**. As the velocity vectors $\vec{v}$ show, each point on the disk has a different velocity $\vec{v}$, with points farther from the axis of revolution having a greater speed than points closer to the axis. Yet all points on the disk execute the same number of revolutions per unit time, and so they all have the same rotational velocity ω_ϑ.

The velocity of an object in circular motion constantly changes direction. As **Figure 11.8*a*** shows, even though the initial and final *speeds* of the object are the same (indicated by the equal magnitudes of the velocity vectors), the vectors $\vec{v}_\text{i}$ and $\vec{v}_\text{f}$ point in different directions. The object therefore undergoes a change in velocity $\Delta\vec{v}$, which, by our definition of acceleration, means that the object accelerates.

Objects in circular motion have a nonzero acceleration even if they are moving at constant speed.

Because $\vec{r}$ and $\vec{v}$ are always perpendicular, when $\vec{r}$ rotates through an angle θ, $\vec{v}$ rotates through the same angle, which means the shaded triangles in Figure 11.8*b* are similar. To determine the change in velocity $\Delta\vec{v} = \vec{v}_\text{f} - \vec{v}_\text{i}$, we slide $\vec{v}_\text{i}$ and $\vec{v}_\text{f}$ along their tangents until their tails meet, as shown in Figure 11.8*b*, then reverse $\vec{v}_\text{i}$ (dashed vector) and add the reversed vector to $\vec{v}_\text{f}$.

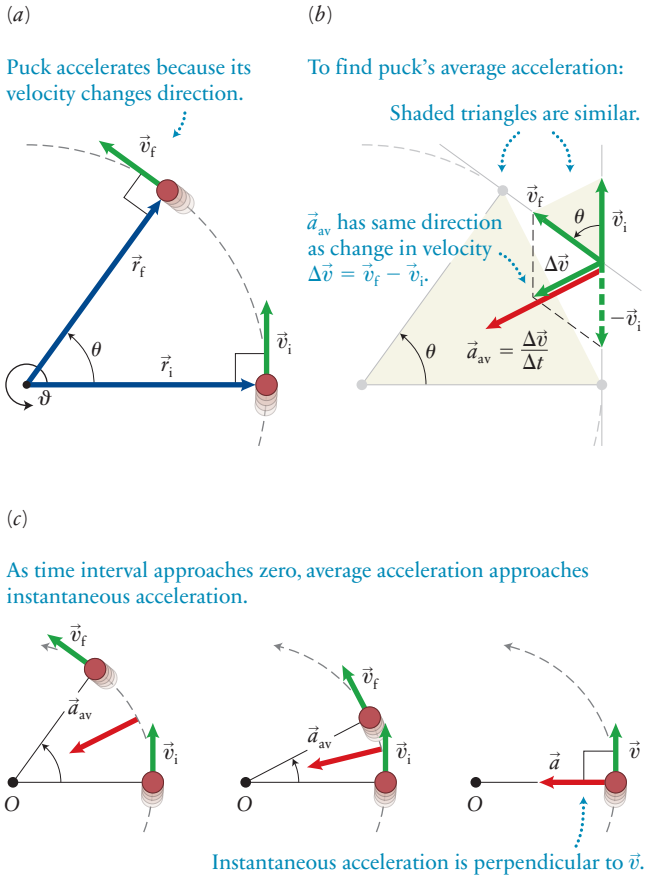
Note that the vector $\Delta\vec{v}$ we obtain in this way points toward the center of the circle. The average acceleration $\vec{a}_\text{av} \equiv \Delta\vec{v}/\Delta t$ therefore also points toward the center. As we let the time interval Δt approach zero, as in Figure 11.8*c*, we see that the (instantaneous) acceleration $\vec{a}$ is perpendicular to the velocity $\vec{v}$. Because the speed of the object is constant, the magnitude of the acceleration is constant too.

An object executing circular motion at constant speed has an acceleration of constant magnitude that is directed toward the center of its circular path.

This acceleration is called the **centripetal acceleration** (*centripetal* is Latin for "center-seeking").

CONCEPTS

Figure 11.8 Average and instantaneous acceleration of an object in circular motion at constant speed.



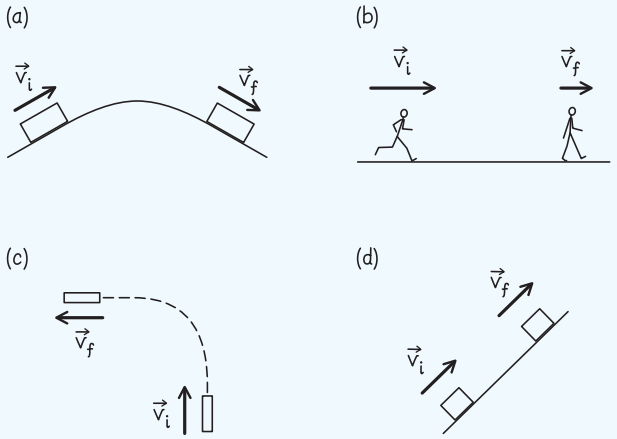
In Chapter 10 we saw that the component of the acceleration parallel to an object's instantaneous velocity causes the speed of the object to change, and the perpendicular component causes the direction of the instantaneous velocity to change. For circular motion at constant speed, there is only a perpendicular component of acceleration (the *centripetal acceleration* we just described), and so only the direction of the velocity changes. Nothing pushes or pulls the object in the direction of motion, and so its speed doesn't change.

EXERCISE 11.2 Accelerations

Determine the direction of the average acceleration in each of the following situations: (*a*) A car goes over the top of a hill at constant speed. (*b*) A runner slows down after crossing a finish line on level ground. (*c*) A cyclist makes a left turn while coasting at constant speed on a horizontal road. (*d*) A roller-coaster car is pulled up a straight incline at constant speed.

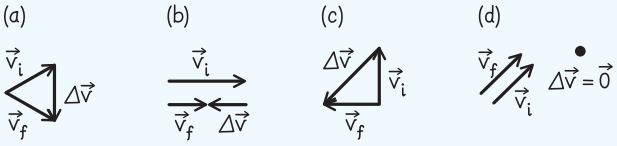
SOLUTION For each situation I make a before-and-after sketch showing the initial and final velocities (**Figure 11.9**). The acceleration is nonzero if the direction of the velocity or the magnitude of the velocity changes.

Figure 11.9



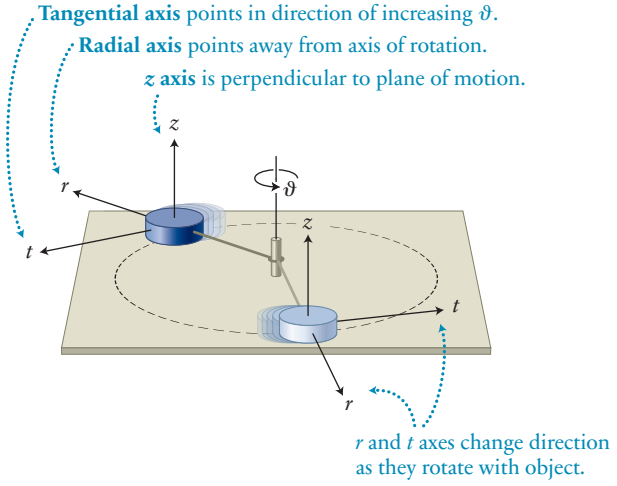
The average acceleration points in the same direction as the change in velocity $\Delta\vec{v}$. To determine the direction of this vector, I draw the vectors $\vec{v}_i$ and $\vec{v}_f$ with their tails together. The change in velocity then points from the tip of $\vec{v}_i$ to the tip of $\vec{v}_f$, as shown in **Figure 11.10**. For each situation, the direction of the average acceleration is given by the direction of $\Delta\vec{v}$. ✔

Figure 11.10



To describe circular motion it is convenient to choose a coordinate system that rotates around with the object under consideration, as illustrated in **Figure 11.11**. The *radial axis*, denoted by the letter *r*, points in the direction of the radius of the circular trajectory, away from

Figure 11.11 Rotating coordinate system used to describe circular motion.



CONCEPTS

the axis of rotation. The *tangential axis*, denoted by the letter t, points in the direction of increasing ϑ, tangent to the trajectory. Because the tangent to a circle is perpendicular to the radius drawn to the point of contact of the tangent, the r and t axes are always perpendicular to each other. Often we'll need a third axis, denoted by the letter z, to describe the direction of forces exerted on the object. This axis is chosen perpendicular to the plane of motion (and therefore perpendicular to both the r and t axes).

With this choice of coordinate system, the velocity of an object executing circular motion at constant speed is always directed along the t axis and the acceleration of the object is always directed inward along the r axis.

11.1 Suppose the object in Figure 11.8 is in *accelerated* circular motion, so that $|\vec{v}_f| > |\vec{v}_i|$. In which direction does the object's average acceleration point?

11.2 Forces and circular motion

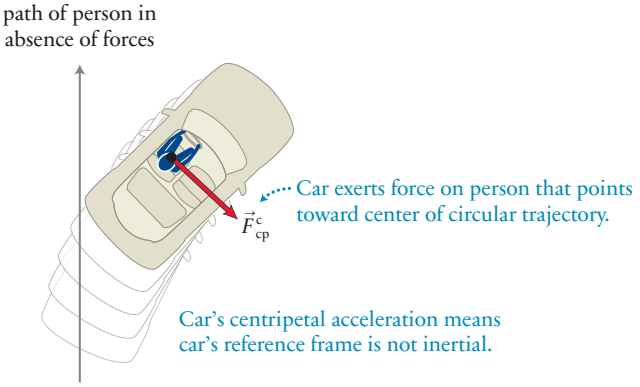
The centripetal acceleration of an object in circular motion at constant speed tells us that the vector sum of the forces exerted on the object must be directed toward the center of the circle, continuously adjusting the object's direction. Without this inward-pointing vector sum of forces, the object would move in a straight line.

An object that executes circular motion at constant speed is subject to a force (or vector sum of forces) of constant magnitude directed toward the center of the circular trajectory.

But wait! Aren't you pushed *outward* when a car makes a sharp turn? Say you are in a car that takes a right turn as illustrated in **Figure 11.12**. Even though the center of the arc describing the car's trajectory is on the right, there is no denying that your left shoulder is pressed against the left door! Don't forget, however, that because the direction of the car's velocity is changing, it has a nonzero acceleration, and so the reference frame of the car is noninertial.

In Chapter 6 we discussed how things are different in noninertial reference frames. Riding in a car that is traveling on a straight road when it makes an emergency stop, you feel as if you are being pushed forward. There is no force pushing you forward, however. Instead, the "push" you think you feel is nothing more than your body obeying the law of inertia and continuing at its original velocity. As your body tries to continue, it collides with the seat belt because the belt, being part of the car, begins to slow down. You exert a forward force on the belt, but it exerts a rearward force on you. Likewise, when a car initially traveling

Figure 11.12 As you round a curve in a car, are you pushed inward or outward?



straight ahead makes a sharp right turn, your body has a tendency to continue in a straight line. As the car begins to curve rightward, the force exerted by the left door on you forces your body into the turn. So even though you feel you are being pushed outward, an inward contact force $\vec{F}^{c}_{cp}$ is exerted by the car on you, as illustrated in Figure 11.12. The feeling of being pushed outward (leftward) arises only from the noninertial nature of the car's reference frame; in reality, your body simply tends to continue moving in a straight line. The moral is:

Avoid analyzing forces from a rotating frame of reference because such a frame is accelerating and therefore noninertial.

The force required to make an object go around a curve can be a push, a pull, or a combination of the two. For example, the puck in Figure 11.2*b* is pulled inward by the string; in Figure 11.12, the person is pushed inward by the car.

The magnitude of the force required to make an object move in circular motion at constant speed depends on the speed of the object and the radius of the trajectory. To see how this force depends on speed, refer to Figure 11.8*b*. By mentally increasing the magnitude of the vectors $\vec{v}_i$ and $\vec{v}_f$ (which is equivalent to increasing the speed of the revolving object), you see how the vectors $\Delta\vec{v}$ and $\vec{a}$ also get longer. Because $\Sigma\vec{F} = m\vec{a}$, an object traveling at higher speed requires a larger inward force.

To see how the inward force depends on radius, look at **Figure 11.13**, which shows two pucks moving at the same speed along circles of different radii. During a given time interval, the rotational coordinate ϑ of the puck moving along the larger circle changes less than that of the puck moving along the smaller circle. Consequently the puck on the larger circle has a smaller change in velocity and a smaller centripetal acceleration. Because force is directly

CONCEPTS

Figure 11.13 Change in velocity for pucks moving along circles of different radius.

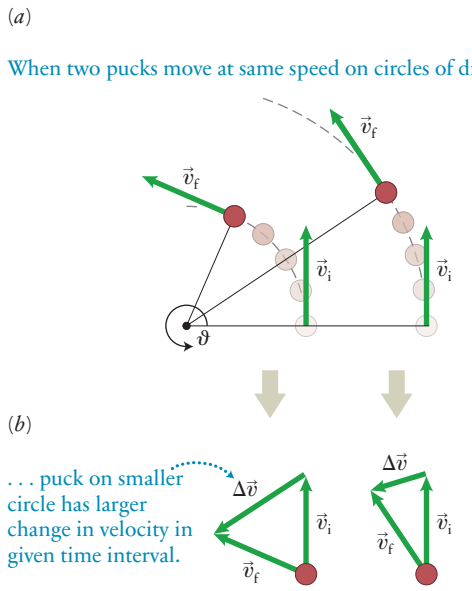
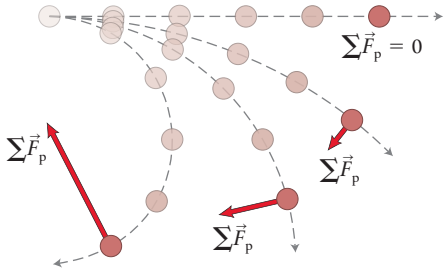


Figure 11.14 The smaller the radius, the sharper the curve and the larger the force required to make the object follow the curve.



proportional to acceleration, we can say that the larger the radius, the smaller the force required to make the object go in a circle (**Figure 11.14**). As the radius approaches infinity, the object goes in a straight line and no force is required to change the direction of its velocity.

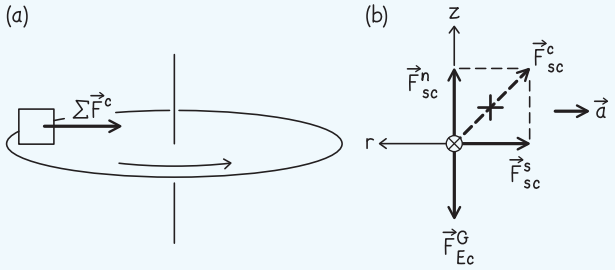
The inward force required to make an object move in circular motion increases with increasing speed and decreases with increasing radius.

EXAMPLE 11.3 Cube on turntable

A cube lies on a turntable initially rotating at constant speed. The rotational speed of the turntable is slowly increased, and at some instant the cube slides off the turntable. Explain why this happens.

1 GETTING STARTED I'm not given much information, so I begin by making a sketch of the situation (**Figure 11.15a**). As the turntable rotates, the cube executes a circular motion. My task is to explain why the cube does not remain on the turntable as the turntable rotates faster.

Figure 11.15



2 DEVISE PLAN Because the cube executes circular motion, it has a centripetal acceleration, and so the vector sum of the forces exerted on it must point toward the center of the turntable. I therefore need to make a free-body diagram that reflects this combination of forces and determine how the forces change as the rotational speed of the turntable increases.

3 EXECUTE PLAN To draw my free-body diagram, I must answer the question What are the forces exerted on the cube? First there is $\vec{F}^G_{Ec}$, the gravitational force exerted by Earth. This force is directed vertically downward and so cannot contribute to a force directed toward the turntable's center. The only other force exerted on the cube is $\vec{F}^c_{sc}$, the contact force exerted by the surface of the turntable. The normal component $\vec{F}^n_{sc}$ of this force must be equal in magnitude to $\vec{F}^G_{Ec}$ because the cube does not accelerate in the vertical direction. The horizontal component, which is the force of static friction $\vec{F}^s_{sc}$, is what forces the cube toward the center of its circular path. I therefore draw the free-body diagram shown in Figure 11.15*b*. Because the vertical component of the acceleration is zero, the forces in that direction add to zero. Thus the vector sum of the forces exerted on the cube equals the force of static friction.

As the rotational speed increases, the magnitude of the centripetal acceleration of the cube also increases. This means that the magnitude of the force of static friction must get larger. At some instant, this force reaches its maximum value and so can no longer increase even though the rotational speed continues to increase. Consequently, the vector sum of the forces exerted on the cube is no longer large enough to give it the centripetal acceleration required for its circular trajectory. When this happens, the distance between the cube and the axis of rotation increases until the cube slides off the edge. ✔

4 EVALUATE RESULT What makes the cube slide off the turntable is its tendency to continue in a straight line (that is, on a trajectory tangent to its circular trajectory). Up to a certain speed the force of static friction is large enough to overcome this tendency and keep the cube moving in a circle. Once the force of static friction reaches its maximum value, the cube begins to slide.

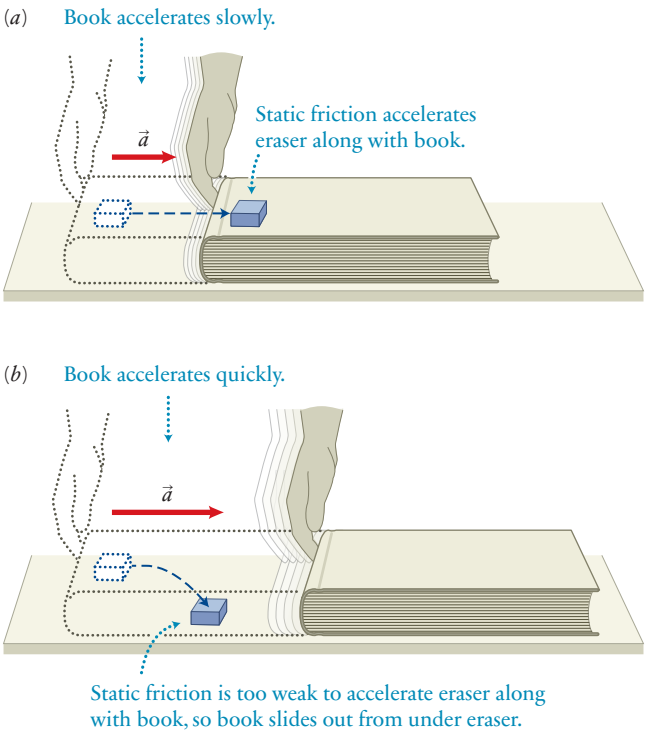
CONCEPTS

It is all too easy to think that there must be an *outward* force pushing the cube off the turntable as the rotational velocity increases in Example 11.3. As in a car rounding a corner, this thought arises because, without realizing it, you tend to view the situation from the rotating frame of the turntable. Because this is an easy mistake to make in situations involving rotation, you must always make sure you choose an inertial frame of reference.

If you can't imagine the cube moving off the turntable without being *pushed* off, try the following: Put some object—an eraser or your keys will do—on top and near the edge of a book as in **Figure 11.16*a*** and then gently, without touching the object, push the book along the surface of a table. The object moves along because it is being pushed forward by the force of static friction exerted by the book on it. If you give the book a hard shove, however, the object lags behind and falls off the rear edge of the book. Did it get pushed off—by, say, air resistance? Someone moving along with the book might claim it did—just as someone in an accelerating car feels "pushed backward"—but such a person is not in an inertial reference frame. (Also, this experiment would work even in a vacuum.) From the point of view of someone in the Earth reference frame, the book got yanked from under the object. The force of static friction exerted by the book on the object is not large enough to give the object the same acceleration as the book, and so the object falls off the rear edge without anything physically pushing it off. If you mark its starting position, you will notice that the object always ends up in front of its starting position, as shown in Figure 11.16*b*. This observation tells you that, even though it falls off the rear edge of the book, the object must have been pushed in the *forward* direction. It falls off the book not because it has been pushed off, but because the forward push exerted by the book on the object is not large enough for the object to keep up with the accelerating book.

The same is true for the cube on the turntable: It falls off the edge not because it is pushed off but because the force of static friction between turntable and cube is not large enough to force the cube to go around in a circle.

Figure 11.16 The eraser in (*b*) falls off the book because of inertia, not because of a force "pushing" it off the book.



11.2 Suppose I have two cubes on a turntable at equal distances from the axis of rotation. The inertia of cube 1 is twice that of cube 2. Do both cubes begin sliding at the same instant if I slowly increase the rotational velocity?

Figure 11.17 shows free-body diagrams for three objects moving in circular motion at constant speed. In each case the vector sum of the forces exerted on the object points toward the center of the circular trajectory.

11.3 (*a*) Does a bicycle always have to lean into a curve as illustrated in Figure 11.17*a*? (*b*) The rope holding the bucket in Figure 11.17*b* makes a small angle with the horizontal. Is it possible to swing the bucket around so that the rope is exactly horizontal?

11.3 Rotational inertia

Having established the principles of circular motion at constant speed, let us now examine *changes* in rotational velocity. Imagine two identical pucks A and B on a horizontal air table (**Figure 11.18**). Puck B is initially at rest and fastened to a vertical post in the middle of the table by a short string that constrains the puck's motion to a circle of radius r_B. Suppose we give puck A a shove so that it strikes stationary puck B at speed v, as illustrated in Figure 11.18*a*. The two pucks collide elastically, and so immediately after the collision puck A is at rest and puck B has speed v. Puck B then begins to move around the post in circular motion at constant speed v. If the change in its rotational coordinate during a time interval Δt is $\Delta\vartheta_B$, then its rotational velocity after the collision is $\omega_{B\vartheta} = \Delta\vartheta_B/\Delta t$.

Now suppose we shove puck A at the same initial speed v toward a third identical puck C initially at rest and attached to a string of length $r_C > r_B$, as in

CONCEPTS

Figure 11.17 Free-body diagrams for three objects in circular motion at constant speed.

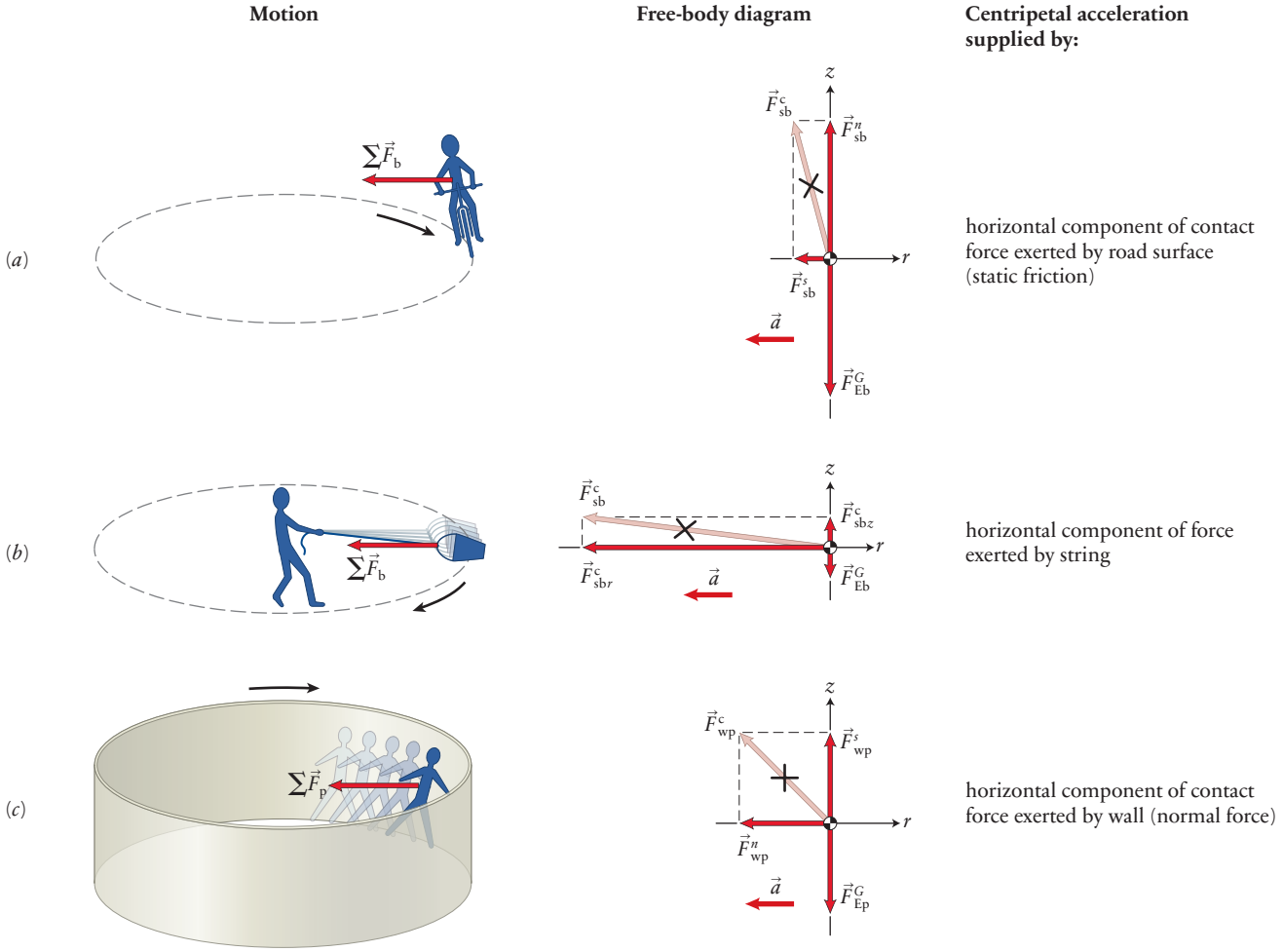


Figure 11.18 When a moving puck strikes a stationary puck attached to a string, the rotational velocity of the puck on the string depends on string length.

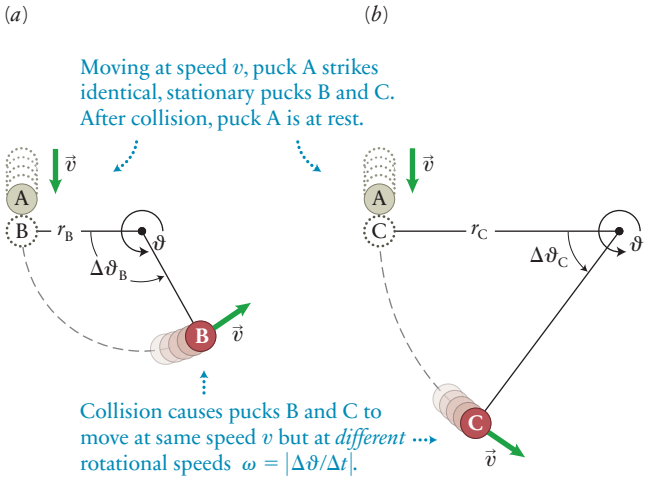


Figure 11.18*b*. As before, puck A comes to rest after the collision and puck C moves away with speed v. Because both pucks B and C have the same speed v after the collision, they travel the same arc length in the same time interval Δt. However, their *rotational* speeds are different because the radii of their trajectories are different. The magnitude of the change $\Delta\vartheta_C$ in the rotational coordinate of C during the time interval Δt is smaller than that of B, and C's rotational speed, $\omega_C = |\Delta\vartheta_C/\Delta t|$, is therefore smaller than that of B.

If puck C is to attain the same final rotational velocity as B, it needs to be struck by a considerably faster moving puck A. Put differently, it is more difficult to change the rotational velocity of C than it is to change that of B. Note that C and B have the same inertia. If it weren't for the presence of the strings of different lengths, we would not be able to distinguish them: Once hit by A, both C and B would move off in a

CONCEPTS

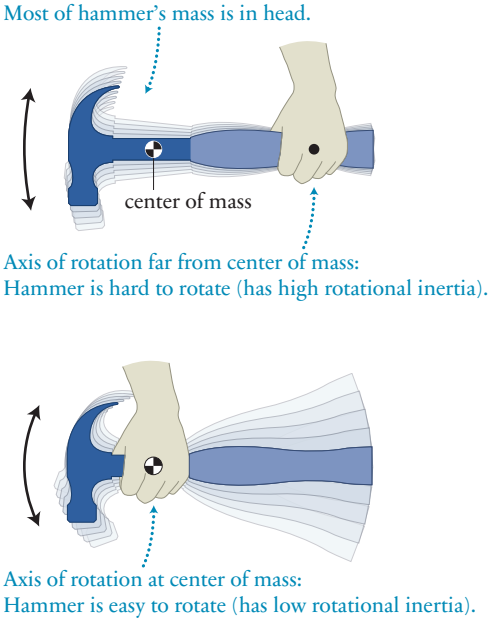
straight line with the same speed v. Apparently an object's tendency to resist a change in rotational velocity, its **rotational inertia,** is not given by simply the object's inertia—the object's tendency to resist a change in velocity as we defined it in Chapter 4. (The inertia surely plays a role: If I increase the inertia of puck B, its rotational speed after the collision is lower and so is its change in rotational velocity.) That the longer string on C causes C's rotational velocity to change less than B's suggests that an object's rotational inertia depends on its location relative to the axis of rotation. The greater the distance between the bulk of an object's material and the axis of rotation, the larger its rotational inertia. You can easily verify this using a hammer (**Figure 11.19**). If you hold the hammer near its head, it is easily rotated. If you hold the hammer at the handle, as you would normally, it is more difficult to rotate.

There are many examples of the effect of rotational inertia. Skaters reduce their rotational inertia, and therefore spin more easily, by holding their arms close to their bodies; you run faster with your legs bent than with them extended. Conversely, when trying to keep your balance on a beam, you spread out your arms. This action moves more of your inertia away from the rotation axis and so increases your rotational inertia, which in turn increases your resistance to rotating off the beam. Tightrope walkers use a long pole for the same reason: The pole puts more inertia farther from the rotation axis and so increases the walkers' rotational inertia and gives them more time to adjust their balance.

One final important point: Unlike inertia, which has a single value, rotational inertia depends on the location of the rotation axis. An object can have a small rotational inertia about one axis and a large rotational inertia about another axis—just think of the hammer in Figure 11.19. Therefore the value of rotational inertia cannot be determined unless you specify the axis of rotation.

11.4 About which axis is the rotational inertia of a pencil (*a*) largest and (*b*) smallest: (1) a lengthwise axis through the core of the pencil; (2) an axis perpendicular to the pencil's length and passing through its midpoint; (3) an axis perpendicular to the pencil's length and passing through its tip?

Figure 11.19 The rotational inertia of a hammer—its resistance to rotational motion—depends on the location of the axis of rotation.

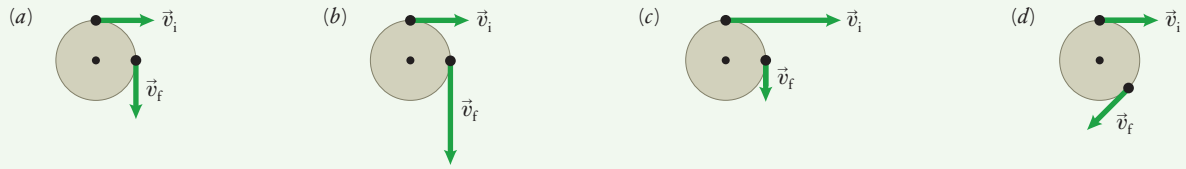


CONCEPTS

Self-quiz

1. Which of the following is in translational equilibrium? (*a*) An object whose center of mass is undergoing circular motion at constant speed. (*b*) A wheel spinning about an axis through its center of mass.
2. **Figure 11.20** shows the velocities $\vec{v}_i$ and $\vec{v}_f$ of a point on the rim of an object at two instants t_i and t_f for four different motions. For each case, determine the change in velocity $\Delta\vec{v}$ and the direction of the average acceleration at a point in time halfway between t_i and t_f.

Figure 11.20

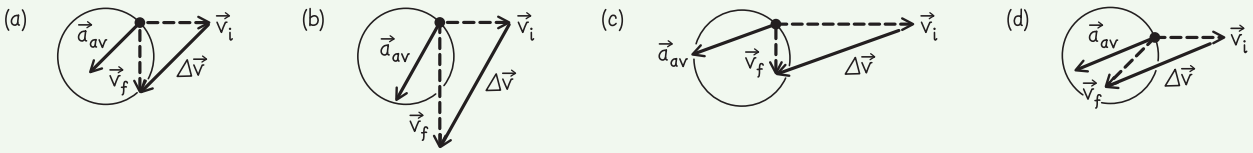


3. Why doesn't the acceleration in question 2 point toward the center of the circle in all cases?
4. Describe the interaction responsible for providing a centripetal acceleration for (*a*) a car rounding a level curve, (*b*) a car rounding a banked curve, (*c*) a coin rotating on a turntable, (*d*) a ball swung through a horizontal circle by a string that sweeps out a cone, (*e*) the Moon orbiting Earth, (*f*) clothes spinning in a dryer, (*g*) a marble rolling along the inside of a horizontal hoop, and (*h*) a ball on a string rolling in a horizontal circle on a surface.
5. A ball attached to a string (the far end of which is fixed) rolls in a horizontal circle. Under which conditions is the string more likely to break: (*a*) when the speed of the ball is increased for a given radius or (*b*) when the length of the string is increased for a given speed?

Answers

1. (*a*) An object undergoing circular motion at constant speed is accelerated because the direction of its center-of-mass velocity changes continuously. If it is accelerated, the vector sum of the forces exerted on the object is not zero, and the object cannot be in translational equilibrium. (*b*) If the center of mass of a spinning wheel is fixed, $a_{cm} = 0$ and so the object is in translational equilibrium.
2. For each case, $\Delta\vec{v} = \vec{v}_f - \vec{v}_i$ and the average acceleration is parallel to $\Delta\vec{v}$ (**Figure 11.21**).

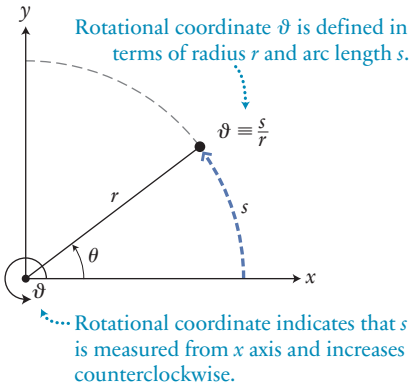
Figure 11.21



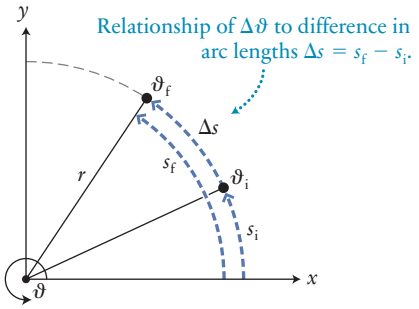
3. Notice that as the object travels around the circle in parts *b* and *c*, its speed changes. To change the speed, there must be a tangential component to the acceleration as well as a centripetal component. The tangential component requires that the acceleration not point directly toward the center.
4. (*a*) Force of static friction exerted by road on car, (*b*) horizontal component of contact force exerted by road on car, (*c*) force of static friction exerted by turntable on coin, (*d*) horizontal component of tensile force exerted by string on ball, (*e*) gravitational force exerted by Earth on the Moon, (*f*) centripetal component of contact force exerted by drum of dryer on clothes, (*g*) contact force exerted by hoop on marble, and (*h*) tensile force exerted by string on ball.
5. In case *a*, for a given radius, a greater speed means the ball travels through a larger angle during a specific time interval. If the angle is larger, the magnitude of $\Delta\vec{v}$ is larger. A larger magnitude of $\Delta\vec{v}$ requires that the acceleration and force also be larger. The force providing the acceleration is due to the tension in the string. Therefore, the greater speed requires more tension and the string is more likely to break. In case *b*, for a given speed, a larger radius means that the ball travels through a smaller angle in a specific time interval. If the angle is smaller, the magnitude of $\Delta\vec{v}$ is smaller. A smaller $\Delta\vec{v}$ requires a smaller acceleration and a smaller force. Therefore, a large radius requires a smaller tension in the string and the string is less likely to break.

Figure 11.22 Relationships between rotational coordinate ϑ, radius r, and arc length s.

(*a*)



(*b*)



11.4 Rotational kinematics

The **rotational coordinate** of an object moving along a circle of radius r is defined as the length of the arc s over which the object has moved divided by the radius:

$$\vartheta \equiv \frac{s}{r}. \tag{11.1}$$

The sign and magnitude of s (and therefore the sign and magnitude of ϑ) depend on our choice of rotational coordinate system. With the choice of rotational coordinate system shown in **Figure 11.22*a***, s is measured along the perimeter of the circle, beginning at the positive x axis and ending at the location of the object, and the counterclockwise direction is taken to be positive. Because the rotational coordinate ϑ is a ratio of two lengths, it is a unitless quantity. In contrast, the polar angle θ swept out by the revolving object can be expressed in *radians* (rad), *degrees* (°), or *revolutions* (rev). The relationship between these units is

$$2\pi \text{ rad} = 360° = 1 \text{ rev}. \tag{11.2}$$

Note that, unlike the units for length, time, and inertia, no physical standard is required to specify the units for polar angle.

Rotational coordinate and polar angle are directly related to each other. After one revolution, the arc length s traveled by the object is equal to the circumference of the circle, and so $s = 2\pi r$. Substituting this value for s in Eq. 11.1 yields $\vartheta = 2\pi$ after one revolution. Knowing from Eq. 11.2 that 2π rad $= 1$ rev, we see that the rotational coordinate corresponds to the *number of radians in the polar angle*: $\vartheta = \theta/(1 \text{ rad})$.* In other words, given the polar angle θ of the object's position, we can obtain its rotational coordinate by expressing the polar angle in radians and dividing by 1 rad.

The change in the rotational coordinate of an object, the object's final rotational coordinate minus its initial rotational coordinate, is

$$\Delta\vartheta \equiv \vartheta_f - \vartheta_i. \tag{11.3}$$

This quantity is positive for a rotation in the direction of increasing ϑ and negative for a rotation in the direction of decreasing ϑ. Combining this expression with Eq. 11.1 allows us to express $\Delta\vartheta$ as a difference in arc lengths:

$$\Delta\vartheta = \vartheta_f - \vartheta_i = \frac{s_f}{r} - \frac{s_i}{r} = \frac{\Delta s}{r}, \tag{11.4}$$

where Δs is defined as the length of the arc between the final and initial locations of the object, as shown in Figure 11.22*b*.

11.5 Starting from a position with rotational coordinate zero, an object moves in the positive ϑ direction at a constant speed of 3.0 m/s along the perimeter of a circle of radius 2.0 m. (*a*) What is the object's rotational coordinate after 1.5 s? (*b*) How long does it take the object to complete one revolution? (*c*) What is its rotational coordinate 1.0 s after passing a point on the circle whose polar angle is 90°?

QUANTITATIVE TOOLS

* The reason for this division by 1 rad is to cancel out the units of radian. The rotational coordinate ϑ is unitless, but the polar angle θ has units of radians. So for an object that has rotated to an angle of 6.1 rad, the rotational coordinate is $(6.1 \text{ rad})/(1 \text{ rad}) = 6.1$; for an object that has rotated to an angle of 265°, the rotational coordinate is $(265°)(2\pi \text{ rad}/360°)/(1 \text{ rad}) = 4.63$.

In analogy to the definition of average velocity, we can define *average rotational velocity* $\omega_{\vartheta,\text{av}}$ as being equal to the change in the rotational coordinate divided by the time interval during which this change takes place:

$$\omega_{\vartheta,\text{av}} \equiv \frac{\Delta\vartheta}{\Delta t}. \tag{11.5}$$

The SI unit of average rotational velocity is s^{-1} (inverse second). The subscript ϑ is there to remind us that the average rotational velocity is a signed quantity that can be positive or negative, depending on the choice of positive direction of rotation for ϑ.

The **rotational velocity** ω_ϑ is obtained by letting the time interval during which the change in the rotational coordinate is measured approach zero:

$$\omega_\vartheta \equiv \lim_{\Delta t\to 0} \frac{\Delta\vartheta}{\Delta t} = \frac{d\vartheta}{dt}. \tag{11.6}$$

The instantaneous rotational velocity is equal to the time derivative of the rotational coordinate just as the instantaneous velocity is equal to the time derivative of position. The **rotational speed** is the absolute value of the rotational velocity:

$$\omega = |\omega_\vartheta|. \tag{11.7}$$

The relationship between instantaneous rotational velocity and instantaneous velocity can be obtained by substituting Eq. 11.4 into Eq. 11.5:

$$\omega_\vartheta = \frac{d\vartheta}{dt} = \frac{1}{r}\frac{ds}{dt}. \tag{11.8}$$

Because the rate at which the arc length s is swept out is the tangential component of the velocity of the object, $v_t = ds/dt$, we have

$$\omega_\vartheta = \frac{v_t}{r} \tag{11.9}$$

or

$$v_t = r\omega_\vartheta. \tag{11.10}$$

Both the tangential component of the velocity v_t and the rotational velocity ω_ϑ are signed quantities that are positive for motion in the direction of increasing ϑ and negative in the direction of decreasing ϑ. We can also express Eq. 11.10 in terms of speeds:

$$v = r\omega. \tag{11.11}$$

In words, the speed of an object in circular motion equals its rotational speed multiplied by the radius of the circular path. As we concluded when we analyzed Figure 11.6, the speed increases with increasing radius.

We can also define a **rotational acceleration** α_ϑ (Greek letter alpha):

$$\alpha_\vartheta \equiv \lim_{\Delta t\to 0} \frac{\Delta\omega_\vartheta}{\Delta t} = \frac{d\omega_\vartheta}{dt} = \frac{d^2\vartheta}{dt^2}. \tag{11.12}$$

The rotational acceleration has SI units s^{-2} (inverse seconds squared). It is a measure of the rate at which the rotational velocity increases.

For *circular motion at constant speed,* the rotational speed $\omega = v/r$ is constant, and so the rotational acceleration is zero. How can rotational acceleration be zero if, as we saw in Section 11.1, circular motion at constant

speed requires a nonzero centripetal acceleration? The answer is that rotational acceleration measures the change in the magnitude of the velocity $\vec{v}$ with which an object travels around a circle, while centripetal acceleration measures the change in the *direction* of $\vec{v}$. Even if the object's speed v is constant, meaning its rotational acceleration is zero, the changing direction of $\vec{v}$ means there is a nonzero centripetal acceleration.

To obtain the magnitude of the centripetal acceleration, remember that if the position vector of a revolving object sweeps out an angle θ, the velocity vector $\vec{v}$ sweeps out the same angle, as illustrated in **Figure 11.23**. This means that the two shaded triangles in Figure 11.23 are similar and so

$$\frac{|\Delta\vec{r}\,|}{r} = \frac{|\Delta\vec{v}\,|}{v} \tag{11.13}$$

or, rearranging terms and dividing left and right by the time interval Δt during which the change in velocity direction took place,

$$\frac{|\Delta\vec{v}\,|}{\Delta t} = \frac{v}{r}\frac{|\Delta\vec{r}\,|}{\Delta t}. \tag{11.14}$$

In the limit for $\Delta t \to 0$, the term $|\Delta\vec{v}\,|/\Delta t$ approaches the magnitude of the (centripetal) acceleration, which we shall denote by a_c, and the term $|\Delta\vec{r}\,|/\Delta t$ approaches the speed v, and so the magnitude of the **centripetal acceleration** is

$$a_c = \frac{v^2}{r} \quad \text{(circular motion)}. \tag{11.15}$$

This is the acceleration required to keep an object moving at speed v along the perimeter of a circle of radius r. As we saw in Section 11.1, the centripetal acceleration of an object executing a circular motion at constant speed is directed radially inward.

Although we derived Eq. 11.15 for circular motion at constant speed, *any* object constrained to move along a circle or part of a circle of radius r is subject to a radially inward acceleration of the magnitude given by Eq. 11.15. Because we defined the radial axis to be directed outward from the rotation axis (Figure 11.11), we can write

$$a_r = -\frac{v^2}{r} \quad \text{(any motion along arc of radius } r\text{)}. \tag{11.16}$$

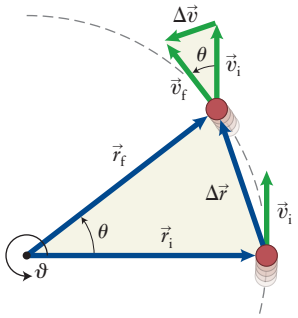
When the object's speed is not constant, we have, in addition to this radial component of the acceleration, a component parallel to the velocity (see Checkpoint 11.1). This component is called the **tangential acceleration** because it is always tangent to the trajectory:

$$a_t = \frac{dv}{dt}. \tag{11.17}$$

The tangential component is positive when the speed increases and negative when it decreases.

Figure 11.24 summarizes a number of important facts about circular motion. Figure 11.24*a* illustrates circular motion at constant speed and Figure 11.24*b* accelerated circular motion. Note that the velocity is always tangential to the trajectory: that is, the radial component of the velocity is

Figure 11.23 Change in the direction of the position $\vec{r}$ and the velocity $\vec{v}$ for an object moving in a circle at constant speed.



QUANTITATIVE TOOLS

always zero. This is true for any type of circular motion, be it at constant speed or accelerated:

$$v_r = 0 \quad \text{(any type of circular motion)}. \tag{11.18}$$

When the circular motion is at constant speed, the radial component of the acceleration is nonzero (Eq. 11.16) and the tangential acceleration is zero:

$$a_t = 0 \quad \text{(circular motion at constant speed)}. \tag{11.19}$$

In that case, the **period** T (the time interval it takes the object to complete one revolution) is

$$T = \frac{2\pi r}{v} = \frac{2\pi}{\omega} \quad \text{(circular motion at constant speed)}, \tag{11.20}$$

where we have used Eq. 11.11 relating the speed v to the rotational speed ω.

If the object speeds up or slows down, both components of the acceleration are nonzero. The acceleration no longer points to the center of the circle, as shown in Figure 11.24*c*, and the magnitude of the acceleration is

$$a = \sqrt{a_r^2 + a_t^2}. \tag{11.21}$$

When the tangential acceleration is nonzero, the rotational acceleration is nonzero too. To relate these two quantities, we substitute Eq. 11.10 for v_t:

$$a_t = \frac{dv_t}{dt} = r\frac{d\omega_\vartheta}{dt}. \tag{11.22}$$

Noting that $d\omega_\vartheta/dt$ is the rotational acceleration α_ϑ (Eq. 11.12), we can write Eq. 11.22 in the form

$$a_t = r\alpha_\vartheta. \tag{11.23}$$

Thus tangential acceleration is related to rotational acceleration in the same way that velocity and rotational velocity are related (Eq. 11.10) and, similarly, arc length and rotational coordinate (Eq. 11.1). For all three we can write (**Table 11.1**):

translational motion quantity $= r \times$ rotational motion quantity.

If a_t is constant (and so the rotational acceleration is constant too), then the arc length traveled and the tangential velocity are given by the kinematics equations for motion at constant acceleration (Eqs. 3.8 and 3.4):

$$s_\text{f} = s_\text{i} + v_{t,\text{i}}\Delta t + \tfrac{1}{2}a_t(\Delta t)^2 \tag{11.24}$$

$$v_{t,\text{f}} = v_{t,\text{i}} + a_t\Delta t. \tag{11.25}$$

Dividing both sides of each of these equations by r and substituting the expressions from Table 11.1, we obtain the equivalent kinematics equations for rotational motion:

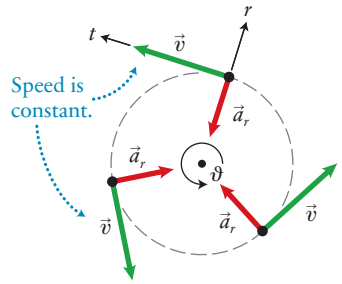
$$\vartheta_\text{f} = \vartheta_\text{i} + \omega_{\vartheta,\text{i}}\Delta t + \tfrac{1}{2}\alpha_\vartheta(\Delta t)^2 \quad \text{(constant rotational acceleration)} \tag{11.26}$$

$$\omega_{\vartheta,\text{f}} = \omega_{\vartheta,i} + \alpha_\vartheta\Delta t \quad \text{(constant rotational acceleration)}. \tag{11.27}$$

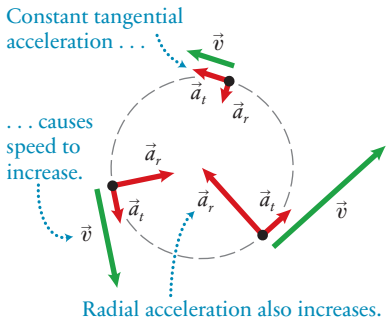
Table 11.2 gives an overview of the correspondence between translational and rotational kinematic quantities and equations.

Figure 11.24 Circular motion and acceleration.

(*a*) Circular motion at constant speed



(*b*) Circular motion at increasing speed



(*c*) Acceleration during circular motion

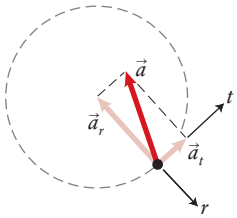


Table 11.1

$s = r\vartheta$	(11.1)
$v_t = r\omega_\vartheta$	(11.10)
$a_t = r\alpha_\vartheta$	(11.23)

Table 11.2 Translational and rotational kinematics

Translational motion (constant acceleration)		Rotational motion (constant rotational acceleration)	
coordinate	x	rotational coordinate	ϑ
x component of displacement	$\Delta x = x_f - x_i$	change in rotational coordinate	$\Delta\vartheta = \vartheta_f - \vartheta_i$
x component of velocity	$v_x = \dfrac{dx}{dt}$	rotational velocity	$\omega_\vartheta = \dfrac{d\vartheta}{dt}$
x component of acceleration	$a_x = \dfrac{dv_x}{dt} = \dfrac{d^2x}{dt^2}$	rotational acceleration	$\alpha_\vartheta = \dfrac{d\omega_\vartheta}{dt} = \dfrac{d^2\vartheta}{dt^2}$
kinematics relationships (constant a_x):		rotational kinematics relationships (constant α_ϑ):	
	$v_{x,f} = v_{x,i} + a_x\Delta t$		$\omega_{\vartheta,f} = \omega_{\vartheta,i} + \alpha_\vartheta\Delta t$
	$x_f = x_i + v_{x,i}\Delta t + \frac{1}{2}a_x(\Delta t)^2$		$\vartheta_f = \vartheta_i + \omega_{\vartheta,i}\Delta t + \frac{1}{2}\alpha_\vartheta(\Delta t)^2$
		radial acceleration	$a_r = -\dfrac{v^2}{r} = -r\omega^2$
		tangential acceleration	$a_t = 0$

EXAMPLE 11.4 Leaning into a curve

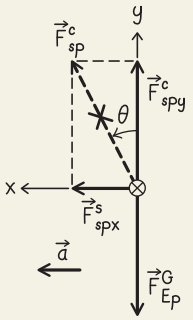
A woman is rollerblading to work and, running late, rounds a corner at full speed, sharply leaning into the curve (**Figure 11.25**). If, during the turn, she goes along the arc of a circle of radius 4.5 m at a constant speed of 5.0 m/s, what angle θ must her body make with the vertical in order to round the curve without falling?

Figure 11.25 Example 11.4.



❶ **GETTING STARTED** As she rounds the circular arc at constant speed, the woman executes circular motion at constant speed. She must therefore undergo a centripetal acceleration as a result of the forces exerted on her. To help me determine the forces exerted on her, I draw a free-body diagram (**Figure 11.26**). The forces exerted on the rollerblader are the gravitational force $\vec{F}^G_{Ep}$ and a contact force $\vec{F}^c_{sp}$ exerted by the surface of the road. Now I see why she must lean into the turn: When she stands straight, the contact force is

Figure 11.26



directed straight up, but as she leans, this force develops a component that pushes her toward the center of the circular arc and provides the necessary centripetal acceleration. I indicate the direction of this centripetal acceleration in my drawing and choose a set of axes—the x axis in the direction of the centripetal acceleration and the y axis upward. (Note that the resulting free-body diagram is similar to that for the bicyclist in Figure 11.17*a*.) I must determine the angle θ that $\vec{F}^c_{sp}$ makes with the vertical.

❷ **DEVISE PLAN** From my free-body diagram, I can draw two conclusions. First, the forces in the y direction must add to zero: $\Sigma F_y = 0$. Second, the x component of the contact force provides the centripetal acceleration given by Eq. 11.15. This gives me two equations from which I should be able to determine θ.

❸ **EXECUTE PLAN** Substituting the centripetal acceleration from Eq. 11.15 into the equation of motion in the x direction, I get

$$\Sigma F_x = ma_x = m(+a_c) = +m\frac{v^2}{r}, \qquad (1)$$

where m is the inertia of the rollerblader. (The rollerblader's inertia is not given, but I hope it will drop out and I won't need it.) From my diagram, I see that

$$\Sigma F_x = F^c_{spx} = F^c_{sp}\sin\theta. \qquad (2)$$

QUANTITATIVE TOOLS

In the y direction I have $F^c_{spy} = F^c_{sp}\cos\theta$ and $F^G_{Epy} = -mg$. The equation of motion in the y direction then yields

$$\Sigma F_y = F^c_{spy} + F^G_{Epy} = F^c_{sp}\cos\theta - mg = 0.$$

Solving this equation for F^c_{sp} and substituting the result into Eq. 2, I get

$$\Sigma F_x = F^c_{sp}\sin\theta = \left(\frac{mg}{\cos\theta}\right)\sin\theta = mg\tan\theta.$$

Substituting this result into Eq. 1 then yields

$$mg\tan\theta = m\frac{v^2}{r}$$

$$\tan\theta = \frac{v^2}{gr} = \frac{(5.0\text{ m/s})^2}{(9.8\text{ m/s}^2)(4.5\text{ m})} = 0.57.$$

This gives an angle $\theta = \tan^{-1}(0.57) = 0.52$ rad, or about 30°. ✔

❹ **EVALUATE RESULT** The rollerblader's angle in Figure 11.25 is about 30°, and so my answer appears to be reasonable.

11.6 Estimate the maximum speed with which you could take the curve in Example 11.4. [Hint: What is the coefficient of static friction?]

11.5 Angular momentum

Let us now return to the experiment discussed in Section 11.3: A stationary puck C fastened to a string of length r is struck by an identical puck A initially moving at speed v. For simplicity we'll ignore the pucks' extent and treat them as particles. After the elastic collision, the puck C moves in circular motion at constant speed v. Calling this puck's inertia m, we can say that its kinetic energy is

$$K = \tfrac{1}{2}mv^2. \tag{11.28}$$

Substituting the expression given by Eq. 11.10 for v, we can rewrite this kinetic energy in the form

$$K = \tfrac{1}{2}mv^2 = \tfrac{1}{2}m(r\omega)^2 = \tfrac{1}{2}(mr^2)\omega^2 \tag{11.29}$$

or, defining the term in parentheses as the **rotational inertia** I of puck C about the axis of rotation

$$I \equiv mr^2 \quad \text{(particle)}, \tag{11.30}$$

we get

$$K_{rot} = \tfrac{1}{2}I\omega^2. \tag{11.31}$$

The rotational inertia of an object about an axis opposes a change in rotational motion about that axis just as inertia opposes a change in translational motion. The SI unit of rotational inertia is $\text{kg}\cdot\text{m}^2$. Because the SI unit of ω is s^{-1}, we get, from Eq. 11.31 SI units of $\text{kg}\cdot\text{m}^2/\text{s}^2$ for the kinetic energy, which is the joule.

Note the remarkable similarity between Eqs. 11.31 and 11.28—inertia m and speed v of Eq. 11.28 are replaced by their rotational equivalents, rotational inertia I and rotational speed ω. The main advantage of writing the kinetic energy in the form of Eq. 11.31 will become clear when we deal with rotating extended objects in Section 11.6: The rotational velocity is the same in all parts of a rotating object, whereas the speed depends on the distance to the axis of rotation (recall Figure 11.7). Because the kinetic energy in Eq. 11.31 is due to rotational motion, it is called the **rotational kinetic energy** of the object.

11.7 Suppose that in Figure 11.18 $r_C = 2r_B$. (*a*) What is the ratio of the rotational velocity of B to that of C? (*b*) What is the ratio of the rotational inertia of B to that of C?

The answers to this checkpoint might surprise you. That the rotational inertia of B is only one-fourth that of C means that B has only one-*fourth* the tendency to resist a change in its rotational motion. So why, when the same force is exerted, is B's rotational velocity only *twice* as large as C's rather than four times as much?

The reason the rotational velocity of B is not as great as you might expect is that the rotational velocity an object acquires depends on how far from the rotation axis it is struck. Increasing the string's length in Figure 11.18 not only increases the puck's rotational inertia but also increases the effectiveness of the blow from the incoming puck. Doubling r for puck C quadruples C's rotational inertia because of the factor r^2 in Eq. 11.30 but also increases puck A's ability to set C in rotational motion. As a result, B's rotational velocity is only twice that of C rather than four times as much.

Figure 11.27 shows two collisions between a bar that is free to rotate about its right end and a puck. In Figure 11.27*a* the puck strikes the end of the rod; in Figure 11.27*b* the rod is struck closer to the axis. Although the rotational inertia of the rod is fixed and although the puck has the same initial momentum, the rod rotates more slowly when struck closer to the rotation axis. (In the extreme case that it is hit right on the axis, the rod doesn't move at all.) So the incoming puck's ability to set other objects in rotational motion depends not only on its momentum but also on how far from the axis it strikes the other object.

Returning to the collision of Section 11.3, we have seen that substituting ω for v and I for m in the expression $\frac{1}{2}mv^2$ yields the final kinetic energy of the struck puck (C in our case), which is equal to the initial kinetic energy of the incoming puck (A):

$$\tfrac{1}{2}mv^2 \rightarrow \tfrac{1}{2}I\omega^2 = \tfrac{1}{2}(mr^2)\left(\frac{v}{r}\right)^2 = \tfrac{1}{2}mv^2. \qquad (11.32)$$

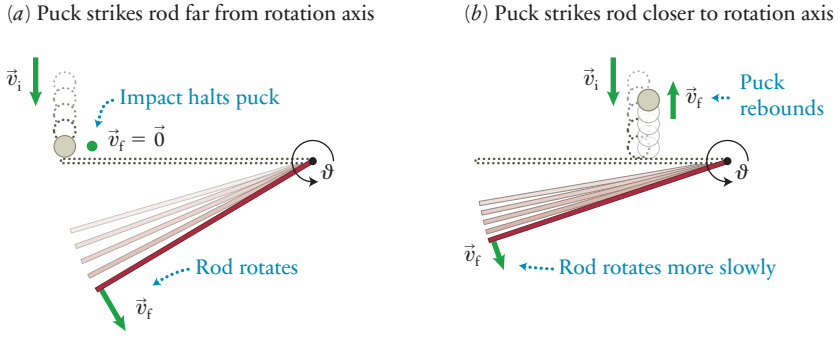
This relationship tells us that the elastic collision causes all the (translational) kinetic energy initially in puck A to be converted to the rotational kinetic energy of puck C. The sum of the kinetic energy of puck A and the rotational kinetic energy of puck C (that is to say, the kinetic energy of the two-puck system) remains constant, as it should.

What happens if we make the same $m \rightarrow I$, $v \rightarrow \omega$ substitutions in mv, the magnitude of the momentum of puck A? Let's try it:

$$mv \rightarrow I\omega = (mr^2)\left(\frac{v}{r}\right) = rmv. \qquad (11.33)$$

The result is different this time: The quantity $I\omega$ is not equal to the magnitude mv of the momentum of the incoming puck A. Instead $I\omega$ is equal to the product of mv and the distance r from A to the axis of rotation at the instant of impact. This "rotational momentum" $I\omega$ thus represents precisely

Figure 11.27 A rod free to rotate about one end is struck by an incoming puck at different distances from its axis of rotation.



QUANTITATIVE TOOLS

what we called the "ability to set in rotational motion" a moment ago: The greater the distance r between the location of impact and the axis of rotation, the larger rmv and the easier it is for puck A to make puck C rotate.

The quantity $I\omega$ plays a fundamental role in physics. The larger the value of $I\omega$ for a moving object, the more easily the object can set another object in rotational motion. Because of the analogy to momentum, which measures an object's capacity to set other objects in motion along a straight line, this quantity is called the **angular momentum** and is denoted by the letter L:

$$L_\vartheta \equiv I\omega_\vartheta. \tag{11.34}$$

The SI units of angular momentum are kg · m^2/s. As before, the subscript ϑ means that the angular momentum, like the rotational velocity, is a signed quantity that can be positive or negative, depending on the choice of the positive direction of rotation for ϑ. Substituting Eqs. 11.10 and 11.30 into Eq. 11.34, we get

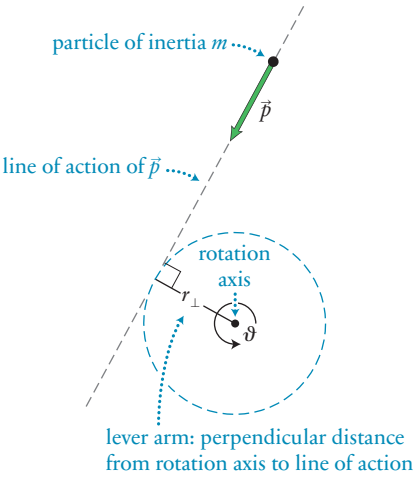
$$L_\vartheta \equiv I\omega_\vartheta = (mr^2)\left(\frac{v_t}{r}\right) = rmv_t. \tag{11.35}$$

An object does not need to rotate or revolve in order to possess a nonzero angular momentum. For example, even though it moves in a straight line, the incoming puck A in Figure 11.18 has a nonzero angular momentum—it has the ability to set another puck in rotational motion. You can use Eqs. 11.34 and 11.35 to calculate the angular momentum of objects in circular motion, but what is the angular momentum of an object that moves along a straight line? The first thing to note is that it doesn't make sense to talk about the angular momentum of an object unless we specify an axis of rotation. **Figure 11.28** shows how to determine the angular momentum for a particle moving along a straight line. For a given axis of rotation (real or imagined), the value of r in Eq. 11.35 is given by the perpendicular distance $r_\perp$ between the axis of rotation and the straight line defined by the momentum of the object (the radius of the circle for which this line, called the *line of action* of the momentum, is a tangent; see Figure 11.28), while the value of v_t is the speed of the object. The distance $r_\perp$ is called the **lever arm distance,** or simply **lever arm,** of the momentum relative to the axis of rotation. The magnitude of the angular momentum of a particle that moves along a straight line is thus

$$L = r_\perp mv \quad \text{(particle)}. \tag{11.36}$$

The sign of the angular momentum is determined by the choice of the positive direction of rotation for ϑ. For example, the rotational coordinate of the particle in Figure 11.28 is increasing as it moves along its straight trajectory, and so the angular momentum of the particle is positive.

Figure 11.28 Relative to an axis of rotation, a particle of inertia m moving along a straight line has an angular momentum of magnitude $r_\perp mv$, where $r_\perp$ is the lever arm distance and v is the speed of the object.



EXERCISE 11.5 Angular momenta

Determine the angular momenta of objects A and B in **Figure 11.29** relative to (*a*) axis 1 and (*b*) axis 2. The inertia of the objects are $m_A = 2.0$ kg and $m_B = 0.50$ kg, and their speeds are $v_A = 0.60$ m/s and $v_B = 1.0$ m/s. The tall side of the rectangle is 2.0 m long and object B travels along the diagonal, which makes an angle of 30° with the long side of the rectangle.

SOLUTION (*a*) Object A: The lever arm distance relative to axis 1 is the long side of the rectangle: $r_\perp = 2.0$ m. The angular momentum is therefore

$$L_{A\vartheta} = +(2.0\text{ m})(2.0\text{ kg})(0.60\text{ m/s}) = +2.4\text{ kg}\cdot\text{m}^2/\text{s}.$$

The plus sign indicates that the rotational coordinate of A relative to axis 1 increases. ✔

Object B: B's line of motion is the diagonal of the rectangle and so goes through axis 1. Therefore $r_\perp = 0$ and thus $L_{B\vartheta} = 0$. ✔

(Continued)

Figure 11.29 Exercise 11.5.

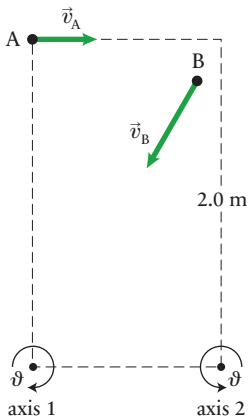
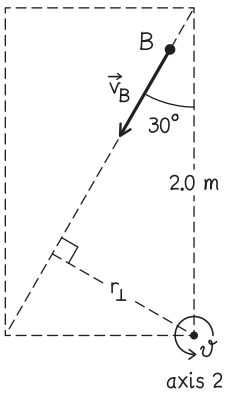


Figure 11.30



(*b*) Object A: The lever arm distance relative to axis 2 is the long side of the rectangle: $r_\perp = 2.0$ m. The angular momentum is therefore $L_{A\vartheta} = -(2.0\text{ m})(2.0\text{ kg})(0.60\text{ m/s}) = -2.4\text{ kg}\cdot\text{m}^2/\text{s}$. The minus sign indicates that the rotational coordinate of A relative to axis 2 decreases. ✔

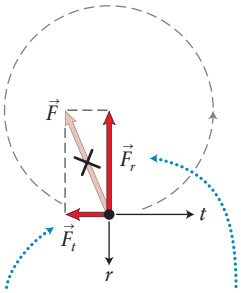
Object B: To determine $r_\perp$ relative to axis 2, I draw a straight line perpendicular to B's line of motion (**Figure 11.30**). Using trigonometry, I see that $\sin 30° = r_\perp/(2.0\text{ m})$, or $r_\perp = (0.5)(2.0\text{ m}) = 1.0$ m, and so $L_{B\vartheta} = +(1.0\text{ m})\ (0.50\text{ kg})(1.0\text{ m/s}) = +0.50\text{ kg}\cdot\text{m}^2/\text{s}$. The plus sign indicates that the rotational coordinate of B relative to axis 2 increases. ✔

Going back to the experiment we discussed in Section 11.3, if we substitute the values of m, $r_\perp$, and v for puck A in Figure 11.18 just before impact, we obtain a value for the angular momentum of puck A just before the impact that is the same as that of puck B after the impact:

$$L_{A\vartheta,i} = L_{B\vartheta,f} = +r_B mv. \tag{11.37}$$

Because puck A comes to rest after the impact, the change in its angular momentum is $\Delta L_{A\vartheta} = L_{A\vartheta,f} - L_{A\vartheta,i} = 0 - r_B mv = -r_B mv$. The change in the angular momentum of puck B is $\Delta L_{B\vartheta} = L_{B\vartheta,f} - L_{B\vartheta,i} = +r_B mv - 0 = +r_B mv$. As you can readily verify, angular momentum is an extensive quantity (see Section 4.4), and so we can define the angular momentum of the system comprising both pucks as $L_\vartheta = L_{A\vartheta} + L_{B\vartheta}$. Because the change in the angular momentum of one puck is compensated for by the change in the angular momentum of the other puck, the angular momentum of the system does not change in the impact: $\Delta L_\vartheta = 0$.

Figure 11.31 Tangential forces change the angular momentum.



Tangential component changes speed and angular momentum.

Radial component causes centripetal acceleration.

This observation suggests that angular momentum, like momentum, is a conserved quantity. Indeed, if it weren't for friction and air resistance, objects in circular motion would tend to continue their circular motion, just as objects in translational motion tend to continue to move along a straight line (see Section 4.1). The more we reduce friction, the longer a spinning wheel keeps turning. In the absence of friction, it would turn forever and its angular momentum would be constant. Consider, for example, the particle in circular motion shown in **Figure 11.31**. The particle is subject to a force $\vec{F}$ that has both a radial and a tangential component. The radial component keeps the particle moving in circular motion. The tangential component, which is tangent to the circular trajectory and points in the direction opposite the direction of motion, causes the particle to slow down. In the absence of any tangential forces, however, the angular momentum remains constant.

Strictly speaking, the fact that the angular momentum of the colliding pucks in Figure 11.18 remains constant is a consequence of conservation of momentum—the change in the momentum of B is compensated for by a change in the momentum of A. It turns out, however, that conservation of angular momentum is just as fundamental as conservation of momentum and conservation of energy. We can generalize this new conservation law to situations where it cannot be deduced from other conservation principles, which tells us that **conservation of angular momentum** is a fundamental law of physics:

Angular momentum can be transferred from one object to another, but it cannot be created or destroyed.

This means that whenever the angular momentum of an object changes, there must be a compensating change of the opposite sign in the environment.

The angular momentum of an object or system on which no tangential forces are exerted is therefore constant:

$$\Delta L_\vartheta = 0 \quad \text{(no tangential forces)}. \tag{11.38}$$

EXERCISE 11.6 Spinning faster

Divers increase their spin by tucking in their arms and legs (Figure 11.32). Suppose the outstretched body of a diver rotates at 1.2 revolutions per second before he pulls his arms and knees into his chest, reducing his rotational inertia from 9.4 kg·m² to 3.1 kg·m². What is his rotational velocity after he tucks in his arms and legs?

Figure 11.32 Exercise 11.6.



SOLUTION Once the diver is off the board, the only force exerted on him is the gravitational force exerted by Earth. This force does not affect the angular momentum of the diver (dropped objects do not spontaneously start to rotate), and so his angular momentum must remain constant. If his angular momentum before he tucks is $L_{\vartheta,\mathrm{i}}$ and that after is $L_{\vartheta,\mathrm{f}}$, then

$$L_{\vartheta,\mathrm{f}} = L_{\vartheta,\mathrm{i}}$$

and so, from Eq. 11.34,

$$I_{\mathrm{f}}\omega_{\vartheta,\mathrm{f}} = I_{\mathrm{i}}\omega_{\vartheta,\mathrm{i}}$$

$$\omega_{\vartheta,\mathrm{f}} = \frac{I_{\mathrm{i}}}{I_{\mathrm{f}}}\omega_{\vartheta,\mathrm{i}}$$

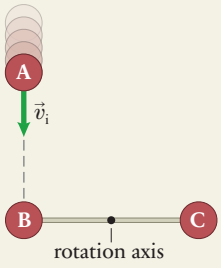
$$= \frac{9.4\ \mathrm{kg\cdot m^2}}{3.1\ \mathrm{kg\cdot m^2}}(1.2\ \mathrm{s^{-1}}) = 3.6\ \mathrm{s^{-1}}. \checkmark$$

11.8 Does the rotational kinetic energy of the diver in Exercise 11.6 change as he pulls his arms in? Explain.

EXAMPLE 11.7 Dumbbell collision

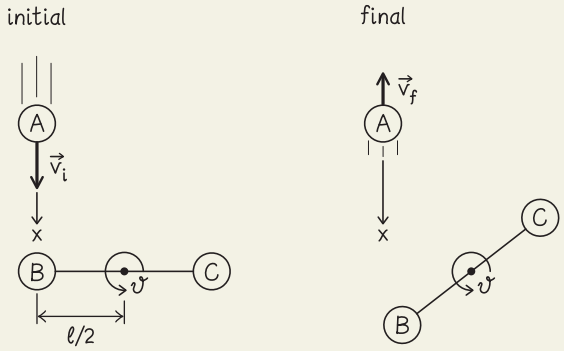
In Figure 11.33, two identical pucks B and C, each of inertia m, are connected by a rod of negligible inertia and length ℓ that is free to rotate about a fixed axis through its center. A third identical puck A, initially moving at speed v_{i}, strikes the combination as shown. After the elastic collision, what are the rotational velocity of the dumbbell and the velocity of puck A?

Figure 11.33 Example 11.7.



1 GETTING STARTED I begin with a two-part sketch (Figure 11.34), choosing an x axis in the direction of A's initial motion and choosing counterclockwise as the positive direction of rotation (this is the direction in which I expect the dumbbell to rotate after the collision). Because A hits B head-on and because the inertia of the dumbbell is twice that of A, I expect A to bounce back and move in the negative x direction after the collision, as my after-collision sketch shows.

Figure 11.34



2 DEVISE PLAN In elastic collisions the kinetic energy of the system remains constant (see Section 5.5). In this collision I need to consider the kinetic energy of puck A and the rotational kinetic energy of the dumbbell. Because there are two unknowns—A's final velocity and the dumbbell's final rotational velocity—I need an additional law to determine both. To this end I apply conservation of angular momentum (Eq. 11.38) to the system comprising puck A and the dumbbell.

(Continued)

❸ **EXECUTE PLAN** The initial kinetic energy of the system is that of puck A, $\frac{1}{2}mv_i^2$. The final kinetic energy is the sum of the (translational) final kinetic energy of A and the rotational kinetic energy of the dumbbell, $K_f = \frac{1}{2}mv_f^2 + \frac{1}{2}I\omega_f^2$, where I is the rotational inertia of the dumbbell. Ignoring the negligible inertia of the rod, I can say that each puck in the dumbbell contributes a rotational inertia $m(\ell/2)^2$ given by Eq. 11.30, so that the rotational inertia of the dumbbell is

$$I = 2m(\ell/2)^2 = \tfrac{1}{2}m\ell^2. \qquad (1)$$

Because the collision is elastic, the final kinetic energy must equal the initial kinetic energy, and so

$$\tfrac{1}{2}mv_i^2 = \tfrac{1}{2}mv_f^2 + \tfrac{1}{2}I\omega_f^2 = \tfrac{1}{2}mv_f^2 + \tfrac{1}{2}(\tfrac{1}{2}m\ell^2)\omega_f^2,$$

where I have substituted for I the expression I obtained in Eq. 1. Dividing both sides by $\frac{1}{2}m$ gives $v_i^2 = v_f^2 + \frac{1}{2}\ell^2\omega_f^2$. Because puck A moves along the x axis, $v_i^2 = v_{x,i}^2$ and $v_f^2 = v_{x,f}^2$. Because $\omega_f^2 = \omega_{\vartheta,f}^2$ I get

$$v_{x,i}^2 = v_{x,f}^2 + \tfrac{1}{2}\ell^2\omega_{\vartheta,f}^2. \qquad (2)$$

Next I turn to conservation of angular momentum. The change in A's angular momentum is, from Eq. 11.36,

$$\Delta L_{A\vartheta} = L_{A\vartheta,f} - L_{A\vartheta,i} = (\ell/2)mv_{x,f} - (\ell/2)mv_{x,i}$$
$$= (\ell/2)m(v_{x,f} - v_{x,i}).$$

The initial angular momentum of the dumbbell $L_{d\vartheta,i}$ about the rotation axis is zero; its final angular momentum about this axis is, from Eq. 11.34, $L_{d\vartheta,f} = I\omega_{\vartheta,f}$. The change in the dumbbell's angular momentum is thus

$$\Delta L_{d\vartheta} = I\omega_{\vartheta,f} - 0 = \tfrac{1}{2}m\ell^2\omega_{\vartheta,f}.$$

Because the system is isolated, its angular momentum doesn't change (Eq. 11.38), and so

$$\Delta L_\vartheta = \Delta L_{A\vartheta} + \Delta L_{d\vartheta}$$
$$= (\ell/2)m(v_{x,f} - v_{x,i}) + \tfrac{1}{2}m\ell^2\omega_{\vartheta,f} = 0$$

or
$$v_{x,i} = v_{x,f} + \ell\omega_{\vartheta,f}. \qquad (3)$$

Solving Eq. 3 for $v_{x,f}$, substituting the result into Eq. 2, and using the quadratic equation to solve for the rotational velocity of the dumbbell, I get

$$\omega_{\vartheta,f} = \frac{4v_{x,i}}{3\ell} = +\frac{4v_i}{3\ell}. \checkmark$$

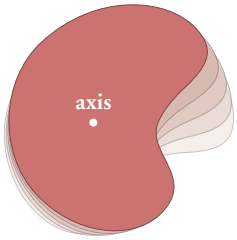
Substituting this result back into Eq. 3, I obtain for the final velocity of puck A $v_{x,f} = -\frac{1}{3}v_{x,i} = -\frac{1}{3}v_i$. ✔

❹ **EVALUATE RESULT** The final rotational velocity is positive, indicating that the dumbbell in Figure 11.34 rotates counterclockwise, in agreement with my drawing. The x component of the final velocity of puck A is negative, indicating that it bounces back, as I expected.

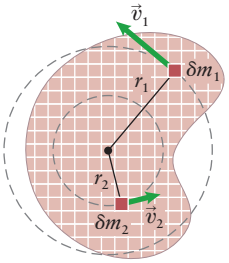
11.9 (*a*) Follow the line of reasoning used in working from Eq. 11.28 to Eq. 11.31 to verify that the rotational inertia of the dumbbell in Example 11.7 is indeed $\frac{1}{2}m\ell^2$. (*b*) Does the momentum of the system remain constant in the collision of Figure 11.33?

Figure 11.35 Procedure for determining the rotational inertia of an extended object.

(*a*) To determine rotational inertia of an extended object . . .



(*b*) . . . divide object into small segments of inertia δm and add up their rotational inertias.



11.6 Rotational inertia of extended objects

We can apply the concept of rotational inertia to extended rigid objects. Consider, for example, the rotating object in **Figure 11.35**. If you imagine breaking down the object into many small segments of equal inertia δm, each segment has a different velocity $\vec{v}$, but all move in circles about the axis at the same rotational speed ω.

The rotational kinetic energy of the rotating object is the sum of the kinetic energies of the segments:

$$K_{rot} = \tfrac{1}{2}\delta m_1 v_1^2 + \tfrac{1}{2}\delta m_2 v_2^2 + \cdots = \sum_n \left(\tfrac{1}{2}\delta m_n v_n^2\right). \qquad (11.39)$$

Because each segment has a different velocity, it is difficult to evaluate this expression. Substituting Eq. 11.11, however, we obtain

$$K_{rot} = \sum_n \left[\tfrac{1}{2}\delta m_n (\omega r_n)^2\right] = \tfrac{1}{2}\left[\sum_n \delta m_n r_n^2\right]\omega^2, \qquad (11.40)$$

where r_n measures the distance from segment n to the rotation axis. Because $\delta m_n r_n^2$ is the rotational inertia I_n of segment n (Eq. 11.30), we can write this expression in the form

$$K_{rot} = \tfrac{1}{2}\left[\sum_n I_n\right]\omega^2 = \tfrac{1}{2}I\omega^2, \qquad (11.41)$$

QUANTITATIVE TOOLS

where I is the rotational inertia of the entire object. So, once we know the object's rotational inertia, it is straightforward to calculate its rotational kinetic energy.

Equations 11.40 and 11.41 show that the rotational inertia of an extended object can be obtained by dividing the object into small segments and computing the sum of the contributions of the small segments:

$$I = \sum_n \delta m_n r_n^2. \tag{11.42}$$

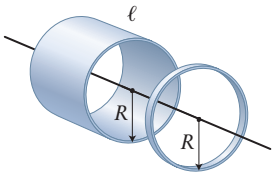
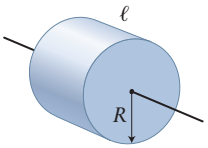
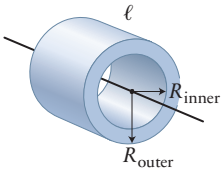
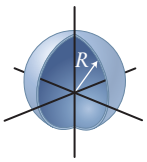
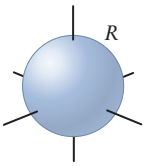
To evaluate this sum for the extended object, we take the limit of this expression as $\delta m_n \to 0$. In this limit, the sum becomes an integral:

$$I = \lim_{\delta m_n \to 0} \sum_n \delta m_n r_n^2 \equiv \int r^2 dm \quad \text{(extended object)}. \tag{11.43}$$

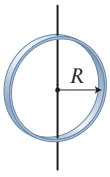
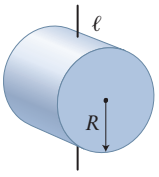
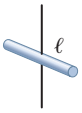
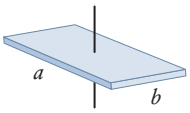
This integral is difficult to evaluate for an arbitrarily shaped object. As the next few examples will show, however, the integral is relatively easy to calculate for objects that exhibit some symmetry and are *uniform* (that is to say, the inertia is uniformly distributed over the volume of the object). **Table 11.3** lists the results of additional calculations.

Table 11.3 Rotational inertia of uniform objects of inertia M about axes through their center of mass

Rotation axes oriented so that object could roll on surface: For these axes, rotational inertia has the form cMR^2, where $c = I/MR^2$ is called the *shape factor*. The farther the object's material from the rotation axis, the larger the shape factor and hence the rotational inertia.

	thin-walled cylinder or hoop	solid cylinder	hollow-core cylinder	thin-walled hollow sphere	solid sphere
Shape					
Rotational inertia	MR^2	$\frac{1}{2}MR^2$	$\frac{1}{2}M(R_{\text{outer}}^2 + R_{\text{inner}}^2)$	$\frac{2}{3}MR^2$	$\frac{2}{5}MR^2$
Shape factor $c = I/MR^2$	1	$\frac{1}{2}$	$\frac{1}{2}\left[1 + \left(\frac{R_{\text{inner}}}{R_{\text{outer}}}\right)^2\right]$	$\frac{2}{3}$	$\frac{2}{5}$

Other axis orientations

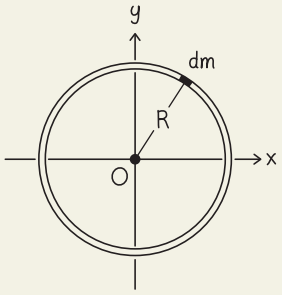
	thin-walled hoop	solid cylinder	thin rod	rectangular plate
Shape				
Rotational inertia	$\frac{1}{2}MR^2$	$\frac{1}{4}MR^2 + \frac{1}{12}M\ell^2$	$\frac{1}{12}M\ell^2$	$\frac{1}{12}M(a^2 + b^2)$

EXAMPLE 11.8 Rotational inertia of a hoop about an axis through its center

Calculate the rotational inertia of a hoop of inertia m and radius R about an axis perpendicular to the plane of the hoop and passing through its center.

❶ **GETTING STARTED** I begin by drawing the hoop and a coordinate system (**Figure 11.36**). Because the axis goes through the center of the hoop, I let the origin be at that location. The axis of rotation is perpendicular to the plane of the drawing and passes through the origin.

Figure 11.36



❷ **DEVISE PLAN** Equation 11.43 gives the rotational inertia of an object as the sum of the contributions from many small segments. If I divide the hoop into infinitesimally small segments each of inertia dm, I see that each segment lies the same distance $r = R$ from the rotation axis (one such segment is shown in Figure 11.36). This means I can pull the constant $r^2 = R^2$ out of the integral in Eq. 11.43, making it easy to calculate.

❸ **EXECUTE PLAN** Substituting $r = R$ in Eq. 11.43, I obtain

$$I = \int r^2\, dm = R^2 \int dm = mR^2. \checkmark$$

❹ **EVALUATE RESULT** This result makes sense because all the material contained in the hoop lies at the same distance R from the rotation axis. Therefore the rotational inertia of the hoop is the same as that of a particle of inertia m located a distance R from the rotation axis, which I know from Eq. 11.30: $I = mR^2$.

For objects in which not all the material is the same distance from the rotation axis, we need to express dm in terms of either Cartesian or polar coordinates in order to evaluate the integral in Eq. 11.43. For uniform one-dimensional objects, we place the object along a positive x axis and introduce a new quantity: *inertia per unit length* $\lambda \equiv dm/dx$. Each small segment dm of the object can then be written as the product of λ and an infinitesimal length dx: $dm = \lambda\, dx$. Equation 11.43 thus becomes, after substituting x for r,

$$I = \lambda \int x^2 dx \quad \text{(uniform one-dimensional object)}, \qquad (11.44)$$

where the integration is carried out over the length of the object.

Likewise, for a uniform two-dimensional object, we convert dm to a surface area by defining an *inertia per unit area* $\sigma \equiv dm/dA$, so that Eq. 11.43 becomes

$$I = \sigma \int r^2 da \quad \text{(uniform two-dimensional object)}, \qquad (11.45)$$

where the integration is carried out over the surface area of the object.

For a uniform three-dimensional object, we introduce an *inertia per unit volume* $\rho \equiv dm/dV$, so that Eq. 11.43 becomes

$$I = \rho \int r^2 dV \quad \text{(uniform three-dimensional object)}, \qquad (11.46)$$

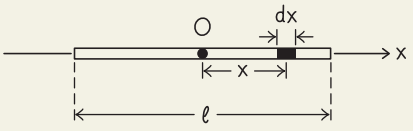
where the integration is carried out over the volume of the object.

EXAMPLE 11.9 Rotational inertia of a rod about an axis through its center

Calculate the rotational inertia of a uniform solid rod of inertia m and length ℓ about an axis perpendicular to the long axis of the rod and passing through its center.

❶ **GETTING STARTED** I begin with a sketch of the rod. For this one-dimensional object, I choose an x axis that lies along the rod's long axis, and because the rotation being analyzed is about a rotation axis located through the rod's center, I choose this point for the origin of my x axis (**Figure 11.37**).

Figure 11.37



❷ **DEVISE PLAN** Because the rod is a uniform one-dimensional object, I can use Eq. 11.44 to calculate its rotational

QUANTITATIVE TOOLS

inertia. First I determine the inertia per unit length λ. Then I carry out the integration from one end of the rod ($x = -\ell/2$) to the other ($x = +\ell/2$).

❸ **EXECUTE PLAN** The inertia per unit length is $\lambda = m/\ell$. Substituting this expression and the integration boundaries into Eq. 11.44, I get

$$I = \lambda \int x^2 dx = \frac{m}{\ell} \int_{-\ell/2}^{+\ell/2} x^2 dx = \frac{m}{\ell}\left[\frac{x^3}{3}\right]_{-\ell/2}^{+\ell/2} = \tfrac{1}{12} m\ell^2. \checkmark$$

❹ **EVALUATE RESULT** If I approximate each half of the rod as a particle located a distance $\ell/4$ from the origin I chose in Figure 11.37, the rotational inertia of the rod would be, from Eq. 11.30, $2(\frac{1}{2}m)(\frac{1}{4}\ell)^2 = \frac{1}{16}m\ell^2$. This is not too far from the value I obtained, so my answer appears to be reasonable.

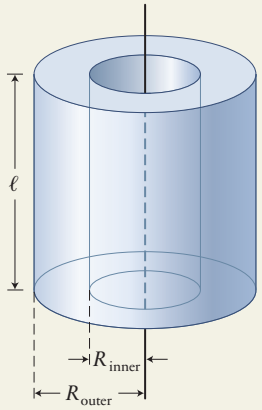


11.10 Calculate the rotational inertia of the rod in Example 11.9 about an axis perpendicular to the long axis of the rod and passing through one end.

EXAMPLE 11.10 Rotational inertia of a hollow-core cylinder

Calculate the rotational inertia of a uniform hollow-core cylinder of inner radius R_{inner}, outer radius R_{outer}, length ℓ, and inertia m about an axis parallel to the cylinder's length and passing through its center, as in **Figure 11.38**.

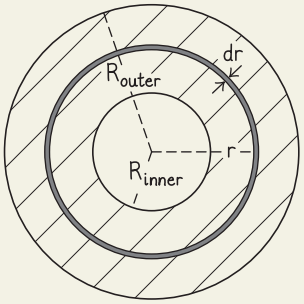
Figure 11.38 Example 11.10.



❶ **GETTING STARTED** As in Example 11.9, I will divide this cylinder into segments and integrate the contributions of all the segments over the volume of the cylinder. There are many ways to divide the cylinder into small segments dm, but I can simplify the integration by exploiting the cylindrical symmetry. Starting at the inner face of the wall, at R_{inner}, and moving toward the outer face, at R_{outer}, I divide the wall into a series of many thin-walled cylindrical shells, each of thickness dr and length ℓ and all concentric with the original cylinder, as shown in my top-down sketch of the cylinder (**Figure 11.39**). Because each shell is infinitely thin, all the material in each shell is the same distance r from the axis of rotation, which simplifies the calculation.

❷ **DEVISE PLAN** The hollow-core cylinder is a uniform, three-dimensional object, so I should use Eq. 11.46. First, I must determine the inertia per unit volume, m/V, for it. Next, I must express the infinitesimal volume dV of each shell in terms of r and dr. Finally, I should carry out the integration for values of r from R_{inner} to R_{outer}.

Figure 11.39



❸ **EXECUTE PLAN** Each thin-walled shell has an inertia $dm = \rho\, dV$, where ρ is inertia per unit volume and dV is the volume of the shell. To determine $\rho = m/V$ for the cylinder, I divide its inertia m by its volume V. The volume of the solid part of the cylinder plus the empty space that forms its core is $\pi R_{\text{outer}}^2 \ell$ and that of the core is $\pi R_{\text{inner}}^2 \ell$. So the volume of the cylinder is $V = \pi(R_{\text{outer}}^2 - R_{\text{inner}}^2)\ell$, and $\rho = m/\pi(R_{\text{outer}}^2 - R_{\text{inner}}^2)\ell$.

Each shell has an outer surface area of $2\pi r\ell$ and thickness dr, and so its volume is $dV = 2\pi r\ell\, dr$. The rotational inertia of the entire cylinder is thus

$$I = \rho \int r^2 dV = \rho \int_{R_{\text{inner}}}^{R_{\text{outer}}} r^2 (2\pi r\ell) dr = 2\pi\rho\ell \int_{R_{\text{inner}}}^{R_{\text{outer}}} r^3 dr.$$

Working out the integral, I get

$$I = 2\pi\rho\ell\left[\frac{r^4}{4}\right]_{R_{\text{inner}}}^{R_{\text{outer}}} = \frac{\pi\rho\ell(R_{\text{outer}}^4 - R_{\text{inner}}^4)}{2}$$

$$= \frac{\pi\ell(R_{\text{outer}}^4 - R_{\text{inner}}^4)}{2} \frac{m}{\pi\ell(R_{\text{outer}}^2 - R_{\text{inner}}^2)}.$$

Factoring $R_{\text{outer}}^4 - R_{\text{inner}}^4 = (R_{\text{outer}}^2 + R_{\text{inner}}^2)(R_{\text{outer}}^2 - R_{\text{inner}}^2)$, I get for the rotational inertia of the hollow-core cylinder

$$I_{\text{hollow-core cylinder}} = \tfrac{1}{2} m(R_{\text{outer}}^2 + R_{\text{inner}}^2). \checkmark$$

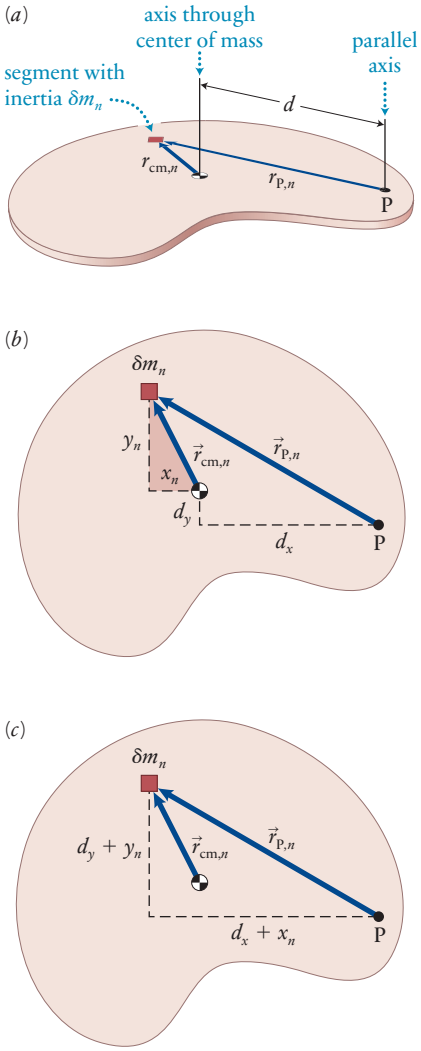
❹ **EVALUATE RESULT** In the limit $R_{\text{inner}} = R_{\text{outer}}$, the cylinder becomes a thin-walled cylindrical shell of radius R_{outer}, and my result becomes $I = mR_{\text{outer}}^2$, as I expect for an object that has all its material a distance R from the axis of rotation.

The result from Example 11.10 allows us to determine the rotational inertia of a solid cylinder by setting R_{inner} equal to zero and R_{outer} equal to the radius R of the cylinder. About the long axis of the cylinder, we have

$$I_{\text{solid cylinder}} = \tfrac{1}{2}mR^2. \tag{11.47}$$

11.11 Note in Example 11.10 that the rotational inertia of a thick cylindrical shell is $\frac{1}{2}m(R_{\text{outer}}^2 + R_{\text{inner}}^2)$ while that of a solid cylinder is $\frac{1}{2}mR_{\text{outer}}^2$. Does the fact that the factor $R_{\text{outer}}^2 + R_{\text{inner}}^2$ is larger than the factor R_{outer}^2 mean that you can increase the rotational inertia of a solid cylinder by drilling a hole through it?

Figure 11.40 The parallel-axis theorem relates the rotational inertia about an axis through the center of mass to the rotational inertia about other parallel axes.



Because rotational inertia depends on the location of the rotation axis, tables such as Table 11.3 may seem to be of limited utility. However, there is a simple relationship between the rotational inertia of an object about an axis through its center of mass and its rotational inertia about any axis parallel to the center-of-mass axis. Consider the object shown in **Figure 11.40**. Let us derive an expression relating the rotational inertia of the object about an axis through its center of mass to the rotational inertia about a parallel axis through an arbitrary point P.

According to Eq. 11.42, the rotational inertia of the object about an axis through its center of mass is

$$I_{\text{cm}} = \sum_n \delta m_n r_{\text{cm},n}^2 = \sum_n \delta m_n (x_n^2 + y_n^2), \tag{11.48}$$

where I have used the Pythagorean theorem for the shaded right triangle in Figure 11.40*b* to substitute for $r_{\text{cm},n}^2$. The rotational inertia about a parallel axis through P a distance d away from the center of mass is

$$I = \sum_n \delta m_n r_{\text{P},n}^2 = \sum_n \delta m_n[(d_x + x_n)^2 + (d_y + y_n)^2], \tag{11.49}$$

where now the Pythagorean substituting comes from the right triangle in Figure 11.40*c*. Working out the right side, I obtain

$$I = \sum_n \delta m_n(d_x^2 + 2d_x x_n + x_n^2 + d_y^2 + 2d_y y_n + y_n^2). \tag{11.50}$$

From the definition of the center of mass (Eq. 6.24), I know that if the center of mass is at the origin, then

$$\sum_n \delta m_n\, x_n = \sum_n \delta m_n y_n = 0, \tag{11.51}$$

and because d_x and d_y are constant, the terms that contain $2d_x x_n$ and $2d_y y_n$ drop out of Eq. 11.50, leaving me with

$$I = \sum_n \delta m_n[(d_x^2 + d_y^2) + (x_n^2 + y_n^2)] = \sum_n \delta m_n d^2 + I_{\text{cm}}. \tag{11.52}$$

The factor d^2 can be pulled out of the summation, and so we see that the rotational inertia about the parallel axis is

$$I = I_{\text{cm}} + md^2. \tag{11.53}$$

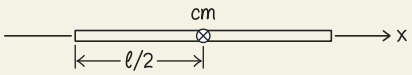
This relationship is called the **parallel-axis theorem.**

EXAMPLE 11.11 Rotational inertia of a rod about an axis through one end

Use the parallel-axis theorem to calculate the rotational inertia of a uniform solid rod of inertia m and length ℓ about an axis perpendicular to the length of the rod and passing through one end.

❶ **GETTING STARTED** I first make a sketch of the rod, showing its center of mass and the location of the rotational axis (**Figure 11.41**). Because I am told to use the parallel-axis theorem, I know I have to work with the rod's center of mass. I know that for a uniform rod, the center of mass coincides with the geometric center, and so I mark that location in my sketch.

Figure 11.41 A solid rod of length ℓ rotating about an axis at one end.



❷ **DEVISE PLAN** In Example 11.9, I determined that the rotational inertia about an axis through the rod's center is $I = \frac{1}{12}m\ell^2$. For a uniform rod, the center of mass coincides with the geometric center, so I can use Eq. 11.53 to determine the rotational inertia about a parallel axis through one end.

❸ **EXECUTE PLAN** The distance between the rotation axis and the center of mass is $d = \frac{1}{2}\ell$, and so, with $I_{cm} = \frac{1}{12}m\ell^2$ from Example 11.9, Eq. 11.53 yields

$$I = I_{cm} + md^2 = \tfrac{1}{12}m\ell^2 + m\left(\frac{\ell}{2}\right)^2 = \tfrac{1}{3}m\ell^2. \checkmark$$

❹ **EVALUATE RESULT** I obtained the same answer in Checkpoint 11.10 by directly working out the integral.



11.12 About what axis is the rotational inertia of an extended object smallest?

Chapter Summary

Rotational kinematics (Sections 11.1, 11.2, 11.4)

Concepts During **rotational motion,** all the particles in an object follow circular paths around the *axis of rotation*.

The **rotational velocity** ω_ϑ of an object is the rate at which the object's rotational coordinate ϑ changes.

The **rotational acceleration** α_ϑ is the rate at which an object's rotational velocity changes.

Quantitative Tools When an object travels a distance s along the circumference of a circle of radius r, the object's **rotational coordinate** ϑ is a unitless quantity defined as s divided by the circle's radius:

$$\vartheta \equiv \frac{s}{r}. \tag{11.1}$$

The arc distance s is measured from the positive x axis. To measure ϑ we need to choose a direction of increasing ϑ and a zero, just as we need to specify a direction of increasing x and an origin to measure position along an axis.

For any rotating object, the rotational velocity and rotational acceleration are

$$\omega_\vartheta = \frac{d\vartheta}{dt} \tag{11.6}$$

$$\alpha_\vartheta = \frac{d\omega_\vartheta}{dt} = \frac{d^2\vartheta}{dt^2}. \tag{11.12}$$

Translational variables for rotating objects (Sections 11.1, 11.4)

Concepts The velocity $\vec{v}$ of an object moving along a circle is always perpendicular to the object's position vector $\vec{r}$ measured from the axis of rotation.

The tangential component v_t of the velocity is tangent to the circle. The radial component v_r of the velocity is zero.

An object moving in a circle has a nonzero acceleration (even if its speed is constant) because the direction of the velocity changes.

An inward force is required to make an object move in a circle, even at constant speed.

Quantitative Tools The tangential and radial components of the velocity of an object moving along a circular path are

$$v_t = r\omega_\vartheta \tag{11.10}$$

$$v_r = 0. \tag{11.18}$$

The radial component of the acceleration is

$$a_r = -\frac{v^2}{r}. \tag{11.16}$$

This radial component is called the **centripetal acceleration** and is directed toward the center of the circle. It can also be written as

$$a_r = -r\omega^2.$$

The tangential component of the acceleration is

$$a_t = r\alpha_\vartheta. \tag{11.23}$$

The magnitude of the acceleration is

$$a = \sqrt{a_r^2 + a_t^2}. \tag{11.21}$$

Constant rotational acceleration (Section 11.4)

Concepts If the tangential acceleration a_t of a rotating object is constant, its rotational acceleration α_ϑ is also constant. In only that case, the rotational kinematics relationships for *constant rotational acceleration* apply.

Quantitative Tools If an object with constant rotational acceleration α_ϑ initially has a rotational coordinate ϑ_i and a rotational velocity $\omega_{\vartheta,\mathrm{i}}$, then after a time interval Δt its rotational coordinate and rotational velocity are

$$\vartheta_\mathrm{f} = \vartheta_\mathrm{i} + \omega_{\vartheta,\mathrm{i}}\Delta\mathrm{t} + \tfrac{1}{2}\alpha_\vartheta(\Delta t)^2 \tag{11.26}$$

$$\omega_{\vartheta,\mathrm{f}} = \omega_{\vartheta,\mathrm{i}} + \alpha_\vartheta\Delta t. \tag{11.27}$$

Rotational inertia (Sections 11.3, 11.5, 11.6)

Concepts **Rotational inertia** is a measure of an object's tendency to resist any change in its rotational velocity. The rotational inertia depends on the inertia of the object and on how that inertia is distributed. The SI units of rotational inertia are kilograms-meters-squared $(\mathrm{kg \cdot m^2})$.

Quantitative Tools The rotational inertia I of a rotating particle of inertia m is

$$I = mr^2, \tag{11.30}$$

where r is the distance from the particle to the rotation axis. For an extended object, the rotational inertia is

$$I = \int r^2 dm. \tag{11.43}$$

The **parallel-axis theorem:** If I_{cm} is the rotational inertia of an object of inertia m about an axis A through the object's center of mass, the rotational inertia I of the object about an axis parallel to A and a distance d away from A is

$$I = I_{cm} + md^2. \tag{11.53}$$

Rotational kinetic energy and angular momentum (Section 11.5)

Concepts **Rotational kinetic energy** is the kinetic energy of an object due to its rotational motion.

Angular momentum L_ϑ is the capacity of an object to make other objects rotate.

A particle can have angular momentum even if it is not rotating.

The law of **conservation of angular momentum** says that angular momentum can be transferred from one object to another but cannot be created or destroyed. The angular momentum of an object or system is constant when no tangential forces are exerted on it.

Quantitative Tools The rotational kinetic energy of an object that has rotational inertia I and rotational speed ω is

$$K_{rot} = \tfrac{1}{2} I\omega^2. \tag{11.31}$$

The angular momentum of an object that has rotational inertia I and rotational velocity ω_ϑ is

$$L_\vartheta = I\omega_\vartheta. \tag{11.34}$$

The angular momentum of a particle of inertia m and speed v about an axis of rotation is

$$L = r_\perp mv, \tag{11.36}$$

where $r_\perp$ is the perpendicular distance from the axis to the line of action of the particle's momentum. The distance $r_\perp$ is called the **lever arm distance** of the momentum relative to the axis.

Questions and Problems

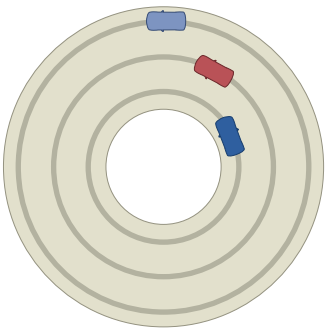
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

11.1 Circular motion at constant speed

1. For a given level of technology, there is a maximum density of "bits" of information that can be stored on a computer's hard-drive disk. The data are stored in a tight spiral, so that "tracks" are essentially circles of increasing radius from inner to outer parts of the disk. As the disk spins at a constant rate, data are read at a constant rate regardless of the radius of the track being read. Which determines the disk capacity: the innermost track or the outermost track? •
2. What is the ratio of the rotational speed of Earth's daily rotation to that of its annual revolution? •
3. When one race car passes another during a race, the passing car usually tries to pass on the inner lane of a curve in the track. Why? •
4. The toy race track in Figure P11.4 has three lanes, and cars cannot change lanes. The centerlines of any two adjacent lanes are 100 mm apart, and the innermost lane has a radius of 1.00 m. (*a*) If each car has the same average speed, compare the time intervals needed for the cars to complete 20 laps. (*b*) If all three cars start beside each other and finish the race at the same instant, how do their average speeds compare? ••

Figure P11.4

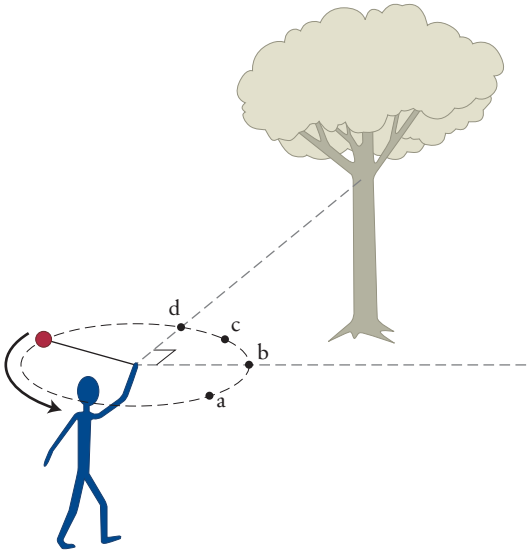


5. Plot, as a function of time, the rotational velocity of your car's wheels as you (*a*) back out of your driveway, (*b*) drive down a residential street, (*c*) turn onto a highway, (*d*) drive on the highway for a while, (*e*) exit the highway and stop at a traffic light, (*f*) drive on a city street, and (*g*) park at your favorite Chinese restaurant. ••
6. Hold a quarter down flat on a table with a finger. Then place a second quarter flat on the table, touching one point on the edge of the first quarter. If you roll the second quarter around the first so that the edges do not slip, describe the revolution and rotation as the second quarter returns to its original location on the table. •••

11.2 Forces and circular motion

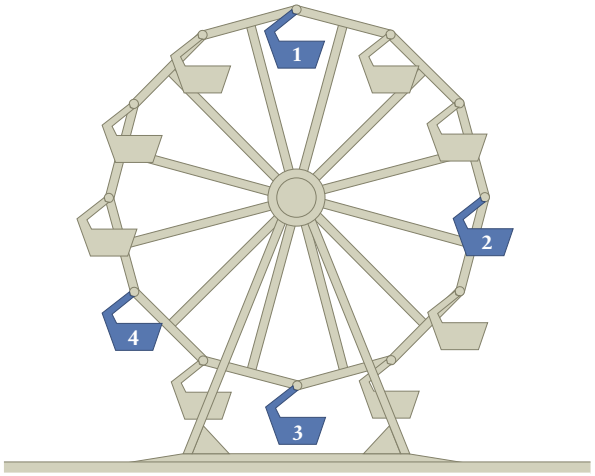
7. Suppose the expression for centripetal acceleration is of the form $a_c = v^p r^q$, where p and q are constants. By analyzing the units in this expression, determine the numerical values of p and q. •
8. Standing in front of a tree, you whirl your sling with a rock in it. If you wish to hit the tree, at what location in the circle in Figure P11.8 (nearest point a, b, c, or d) should you release the rock from the sling? •

Figure P11.8



9. The Ferris wheel in Figure P11.9 is turning at constant speed. Draw and label free-body diagrams showing the forces on passengers at the numbered positions. Be careful to show correct relative lengths on your vector arrows. •

Figure P11.9



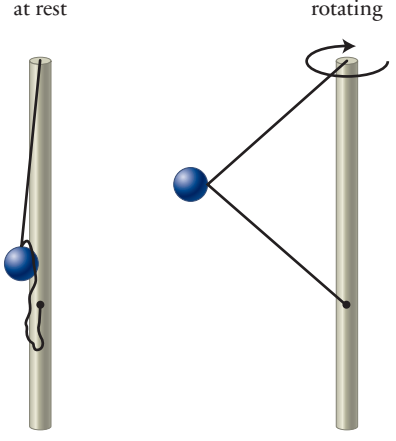
10. To see some of the trouble you can get into with noninertial reference frames, try this experiment. You and a friend stand opposite each other on the edge of a playground merry-go-round and get it rotating. While it is rotating, you attempt to toss a ball to your friend. As seen from above, what does the path of the ball look like (*a*) to an observer standing at the top of a nearby slide, (*b*) to you, and (*c*) to your friend? (*d*) If you did not know you were in a rotating reference frame, how would you explain the path of the ball in part *b*? ••
11. A salad spinner (Figure P11.11) is used to dry lettuce that has been washed. What physics is going on as the spinner does its job? ••

Figure P11.11



12. A race car is negotiating a curve on a banked track. There is a certain speed $v_{critical}$ at which friction is not needed to keep the car on the track in the curve. (This is the speed the car should travel if there is an oil slick on the track, for example.) Draw a free-body diagram showing the forces exerted on the car when it is moving at speed (*a*) $v = v_{critical}$, (*b*) $v > v_{critical}$, and (*c*) $v < v_{critical}$. ••
13. In a certain elaborate roller coaster, the cars are not on tracks but inside a U-shaped track much like a bobsled run. The cars are thus free to swing left and right across the U, climbing the walls as they do so. A student on the ride holds a string with a chunk of metal tied to the free end, with the chunk dangling between his knees as the ride begins. When the car enters a tight left turn, does the chunk swing toward the student's left knee, toward his right knee, or in some other direction? Ignore any friction between cars and track. ••
14. The ball in Figure P11.14 is attached to a vertical rod by two strings of equal strength (able to support much more than one of these balls) and equal length. The strings are

Figure P11.14



light and do not stretch. The rod begins to rotate with its rotational speed increasing slowly, causing the ball to revolve with slowly increasing speed. (*a*) Which string breaks first? (*b*) To justify your answer to part *a*, draw a free-body diagram for the ball, indicating all forces exerted on it and their relative magnitudes. ●●

15. Suppose that you were to tie a heavy metal washer to one end of a string, hold the string by the other end, and whirl it in a horizontal circle above your head. Pay close attention to where your hand would be and to where the washer would be as you speed up the rotation. Ignore air resistance. (*a*) Draw the motion of your hand and the motion of the washer. At a given instant, draw the location of the washer and your hand and the string between them. (*b*) As you speed up the rotation, what force is responsible for the acceleration of the washer? Is there a tangential component to this acceleration? A radial component? (*c*) When the washer achieves some given rotational speed, you cease increasing that speed. Describe the motion of your hand now. ●●●

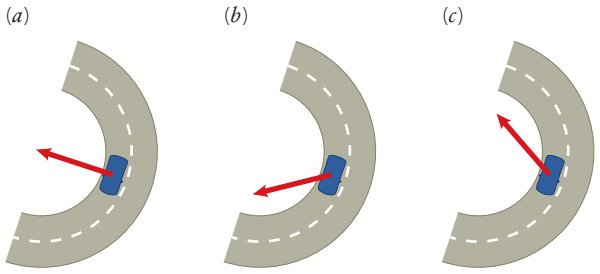
11.3 Rotational inertia

16. Floor lamps usually have a base with large inertia, while the long body and top have much less inertia. If you want to show off by spinning a floor lamp like a baton, where should you grab the lamp? ●
17. A gymnast's backflip is considered more difficult to do in the layout (straight body) position than in the tucked position. Why? ●
18. An ice cube and a rubber ball are both placed at one end of a warm cookie sheet, and the sheet is then tipped up. The ice cube slides down with virtually no friction, and the ball rolls down without slipping. If the ball and the ice cube have the same inertia, which one reaches the bottom first? ●
19. (*a*) Is there a choice of rotation axes for a bowling ball and a baseball, each rotating on its own axis, such that the baseball has a greater rotational inertia? If so, describe the axes. If not, explain why not. (*b*) Repeat for a revolving bowling ball and a revolving baseball. ●●
20. You are designing a uniform thin rectangular door to fill a 1 m wide × 2 m tall opening, and you want it to be easy to open. (*a*) Considering only vertical and horizontal axes, what placement of the axis of rotation would result in the smallest rotational inertia? (*b*) Is such an axis practical? ●●
21. You stand on the edge of a playground carousel while your friends help you get it going. After they stop pushing, you walk toward the center. (*a*) What happens to the rotational inertia of the carousel-you system as you move inward from the edge? (*b*) How does this change in rotational inertia affect the motion of the carousel? ●●

11.4 Rotational kinematics

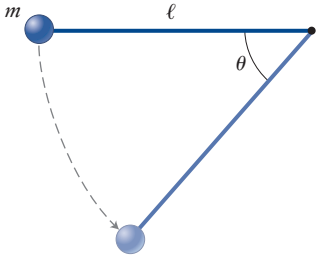
22. A car's engine is turning the crankshaft at 6000 rev/min. What is the rotational speed ω? ●
23. What is the rotational speed of a watch's second hand? Its hour hand? ●
24. If your tires are worn, does your odometer (which is keyed to the turns of the axle) underestimate or overestimate the number of miles you drive? ●
25. Moving at its maximum safe speed, an amusement park carousel takes 12 s to complete a revolution. At the end of the ride, it slows down smoothly, taking 2.5 rev to come to a stop. What is the magnitude of the rotational acceleration of the carousel while it is slowing down? ●●
26. You have a weekend job selecting speed limit signs to put at road curves. The speed limit is determined by the radius of the curve and the bank angle. For a turn of radius 400 m and a 7.0° bank angle, what speed limit should you post so that a car traveling at that speed negotiates the turn successfully even when the road is wet and slick? ●●
27. As you drive clockwise around a turn (as viewed from overhead), you see backed-up traffic ahead and so you slow down. Which force diagram in Figure P11.27 best illustrates your acceleration? ●●

Figure P11.27



28. At the end of a ride, a carnival carousel decelerates smoothly from some initial rotational speed ω to a stop. (*a*) Plot, as a function of time, the speed v of a carousel horse at radius R and of another horse at radius $2R$. (*b*) Plot the magnitude of the centripetal acceleration of the two horses as a function of time. ●●
29. You attach one end of a string of length ℓ to a small ball of inertia m. You attach the string's other end to a pivot that allows free revolution. You hold the ball out to the side with the string taut along a horizontal line, as in Figure P11.29. (*a*) If you release the ball from rest, what is the tension $\mathcal{T}$ in the string as a function of the angle θ swept through? (*b*) What maximum tension should the string be able to sustain if you want it not to break through the ball's entire motion? ●●

Figure P11.29

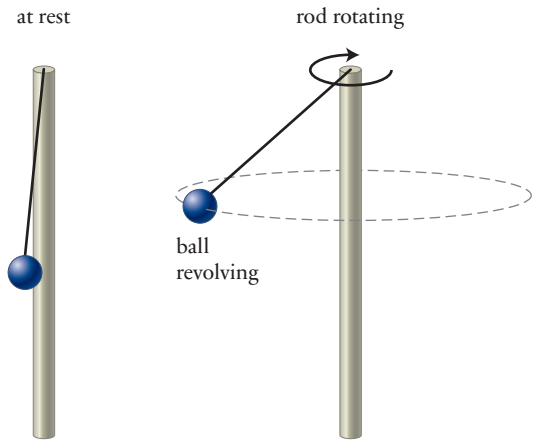


30. A pendulum is made of a bob dangling from a lightweight string of length ℓ. The bob is pulled sideways so that the string makes an angle θ_i with respect to the vertical, then released. As it swings down, what is the rotational speed of the bob as a function of the changing angle θ? ●●

31. Assume that Earth's orbit around the Sun is a perfect circle. Earth's inertia is 5.97×10^{24} kg, the radius of its orbit is 1.50×10^{11} m, and its orbital period is 365.26 days. (*a*) What is the magnitude of its centripetal acceleration around the Sun? (*b*) What are the magnitude and direction of the force necessary to cause this acceleration? ●●

32. A ball is suspended by a string from the top of a vertical rod as shown in Figure P11.32. The rod starts at rest and then begins to rotate with a small constant rotational acceleration. Qualitatively plot as a function of time the magnitude of (*a*) the vertical component and (*b*) the horizontal component of the tensile force exerted on the ball by the string. ●●

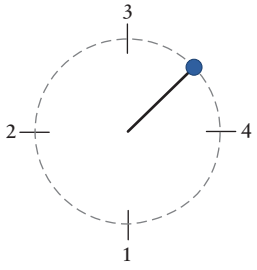
Figure P11.32



33. You hold a small ice cube near the top edge of a hemispherical bowl of radius 100 mm. When you release the cube from rest, what is the magnitude of its acceleration at the instant it reaches the bottom of the bowl? Ignore friction. ●●

34. A 0.100-kg ball attached to a 0.20-m rod of very small inertia moves in a vertical circle in such a way that its speed remains constant. Let $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_3$, and $\mathcal{T}_4$ be the tensions in the rod at positions 1, 2, 3, and 4 in Figure P11.34. (*a*) Rank the tensions in order of magnitude, greatest first. (*b*) What rotational speed is required at position 3 to keep $\mathcal{T}_3$ barely greater than zero? (*c*) At that speed, what is the tension at position 1? ●●

Figure P11.34

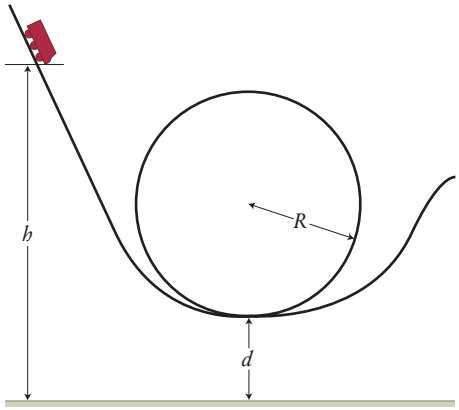


35. An automobile accelerates from rest at $t = 0$ such that its tires undergo a constant rotational acceleration $\alpha = 5.2\ \text{s}^{-2}$. The radius of each tire is 0.38 m. At $t = 10$ s, compute (*a*) the rotational speed ω of the tires, (*b*) the rotational displacement $\Delta\vartheta$ of each tire, (*c*) the speed v of the automobile (assuming the tires stay perfectly round), and (*d*) the distance it has traveled. ●●

36. For the tread on your car tires, which is greater: the tangential acceleration when going from rest to highway speed as quickly as possible or the centripetal acceleration at highway speed? ●●

37. A roller-coaster car initially at a position on the track a height h above the ground begins a downward run on a long, steeply sloping track and then goes into a circular loop-the-loop of radius R whose bottom is a distance d above the ground (Figure P11.37). Ignore friction. (*a*) What is the car's speed when it reaches the bottom of the loop? (*b*) What is the magnitude of the normal force exerted on the car at that instant? (*c*) What is its speed when its position is one-quarter of the way around the loop? (*d*) What is the magnitude of the normal force exerted on it at the one-quarter position? (*e*) What is the car's acceleration at the one-quarter position? ●●●

Figure P11.37

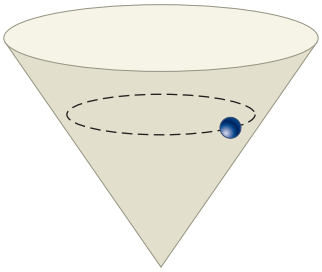


38. (*a*) Consider a circular vertical loop-the-loop on a roller coaster. Show that if a car coasts without power around the loop, the difference between the normal force exerted by the car on a passenger at the top of the loop and the normal force exerted by the car on her at the bottom of the loop is six times the gravitational force exerted on her. (*b*) This difference of $6mg$, referred to as "six gees," is quite a difference for the body to tolerate. To avoid this stress, vertical loops are teardrop-shaped rather than circular, designed so that the centripetal acceleration is constant all around the loop. How must the radius of curvature R change as the car's height h above the ground increases in order to have this constant centripetal acceleration? Express your answer as a function: $R = R(h)$. ●●●

39. A ball is put into a cone and made to move at a constant speed of 3.00 m/s in a horizontal circle of radius 0.500 m (Figure P11.39). (*a*) What is the centripetal component of the ball's acceleration? (*b*) What is the tangential component of its acceleration? (*c*) What force counteracts the force of gravity to keep the ball moving in a horizontal circle? (*d*) Use these insights to determine the height h at which the ball is circling above the bottom of the cone.

[Hint: This is equivalent to determining the angle the cone makes with its vertical axis.] •••

Figure P11.39



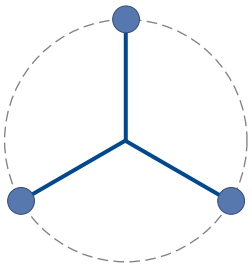
40. A race car driving on a banked track that makes an angle θ with the horizontal rounds a curve for which the radius of curvature is R. (*a*) As described in Problem 12, there is one speed v_{critical} at which friction is not needed to keep the car on the track. What is that speed in terms of θ and R? (*b*) If the coefficient of friction between tires and road is μ, what maximum speed can the car have without going into a skid when taking the curve? •••

11.5 Angular momentum

41. A 0.25-kg ball attached to a string is being spun around on a horizontal surface at a rotational speed of 2.5 rev/s. What is the ball's kinetic energy if the string is 1.0 m long? •

42. Three small, identical 1.0-kg pucks are attached to identical 0.50-m strings, tied together at a common center as shown in Figure P11.42. When whirled in circular motion at angular speed 3.0 s^{-1}, what are (*a*) the rotational kinetic energy and (*b*) the magnitude of the angular momentum about the common center? •

Figure P11.42



43. As you stand on the sidewalk outside an apartment building, a 1.0-kg bird flies toward you at 1.5 m/s along a straight line that passes 2.0 m above your head and runs parallel to the street. A friend is reclining with his head 6.0 m directly above your head, watching from a balcony. (*a*) What magnitude of the angular momentum do you observe the bird to have, using your head as a reference point? (*b*) What magnitude of the angular momentum does your friend observe the bird to have, using his own head as a reference point? •

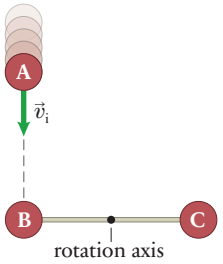
44. Puck A, of inertia m, is attached to one end of a string of length ℓ, and the other end of the string is attached to a pivot so that the puck is free to revolve on a smooth horizontal surface. Puck B, of inertia $8m$, is attached to one end of a string of length $\ell/2$, and the other end of the string is attached to a second pivot so that B is also free to revolve. In each case, the puck is held as far as possible from the pivot so that the string is taut and then given an initial velocity $\vec{v}$ perpendicular to the string. (*a*) How does the magnitude of the angular momentum of puck A about its pivot compare with that of puck B about its pivot? (*b*) How does the rotational kinetic energy of A compare with that of B? ••

45. You have a pail of water with a rope tied to the handle. If you whirl the pail in a vertical circle fast enough, none of the water spills out, even when the pail is upside down. Explain why. ••

46. Consider the roller-coaster car in Problem 37. At what minimum value of the starting height h above the ground must the car begin its journey if it is to remain on the track at the top of the loop? ••

47. In Figure P11.47, two identical pucks B and C, each of inertia m, are connected by a rod of length ℓ and negligible inertia that is free to rotate about its center. Then puck A, of initial speed v_i and inertia $m/2$, comes in and hits B. After the collision, in which no energy is dissipated, what are (*a*) the rotational speed of the dumbbell and (*b*) the velocity $\vec{v}$ of A? ••

Figure P11.47



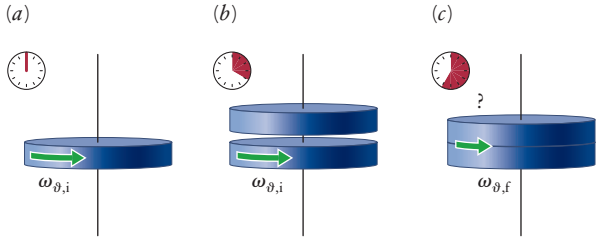
48. A small puck on an air table revolves in a circle with rotational speed ω, held at radius r by a weighted string that passes through a hole in the table. You slowly pull down on the weighted end of the string, decreasing the radius r. Assume that the angular momentum of the puck about the hole remains constant. (*a*) What is the rotational speed when half the string has been taken up? (*b*) What has happened to the speed v during this time interval? ••

49. Two skaters skate toward each other, each moving at 3.3 m/s. Their lines of motion are separated by a perpendicular distance of 2.0 m. Just as they pass each other (still 2.0 m apart), they link hands and spin about their common center of mass. What is the rotational speed of the couple about the center of mass? Treat each skater as a particle, one with an inertia of 75 kg and the other with an inertia of 48 kg. ••

50. A record player turntable initially rotating at $33\frac{1}{3}$ rev/min is braked to a stop at a constant rotational acceleration. The turntable has a rotational inertia of 0.020 kg · m^2. When it is switched off, it slows down to 75% of its initial rotational speed in 5.0 s. (*a*) How long does it take to come to rest? (*b*) How much work has to be done on the turntable to bring it to rest? ••

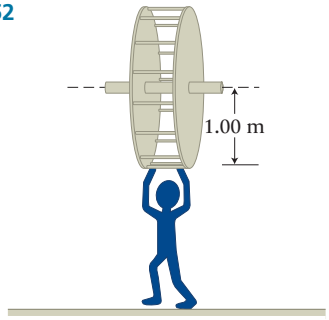
51. A disk of rotational inertia I about the central axis shown in Figure P11.51 is rotating about this axis with initial rotational velocity $\omega_{\vartheta,\mathrm{i}}$ on low-friction bearings. A second identical disk is held at rest a few millimeters directly above the first disk and suddenly dropped. After some slipping, the two discs are observed to have a common rotational speed about the original axis of rotation. What is the magnitude of the final rotational velocity of the combined disks? ••

Figure P11.51



52. A popular ride at some playgrounds consists of two circular disks joined by rungs (Figure P11.52). The center of the unit is attached to a horizontal axle around which the unit rotates. The rotational inertia of the unit is 200 kg · m^2, and its radius is 1.00 m. A 30.0-kg child runs and grabs the bottom rung of the ride when it is initially at rest. Because the child can pull herself up to meet the rung at waist height, treat her as a particle located on the rung. (*a*) If the friction between unit and axle is minimal and can be ignored, what is the minimum speed v the child must have when she grabs the bottom rung in order to make the unit rotate about its axle and carry her over the top of the unit? (*b*) Is this a reasonable speed for a child? •••

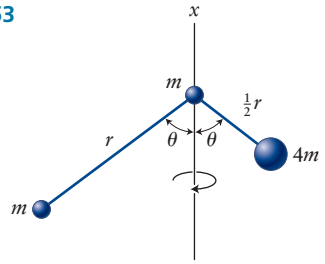
Figure P11.52



11.6 Rotational inertia of extended objects

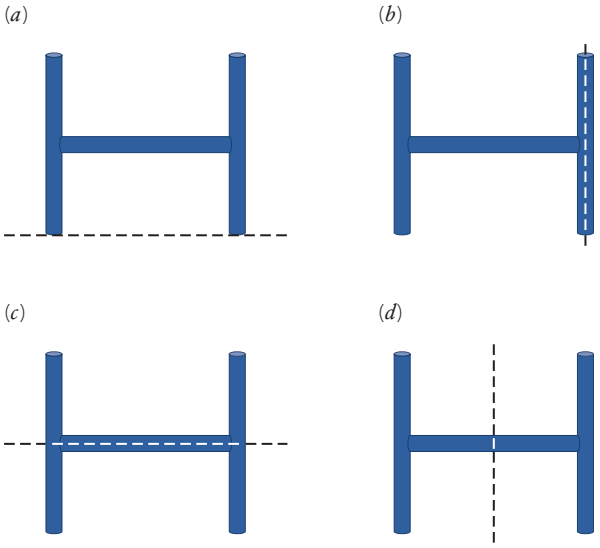
53. What is the rotational inertia about the x axis of the rigid object in Figure P11.53? (Treat the balls as particles.) •

Figure P11.53



54. A cubical wire framework in which each edge is 0.25 m long has a 0.20-kg steel ball at each of its eight corners. What is the rotational inertia of this framework around an axis along one edge? Ignore the inertias of the wires. •
55. What is the formula for the rotational inertia of a thin spherical shell of inertia m and radius R about an axis tangent to any point on its surface? •
56. A deer's leg is very slender near the hoof, with most of the muscle bunched up near the hip joint. How does this work to the deer's advantage? •
57. A 30-kg child running at 1.5 m/s jumps onto a playground merry-go-round that has inertia 190 kg and radius 1.7 m. He is moving tangent to the platform when he jumps, and he lands right on the edge. What is the rotational speed of the merry-go-round and the child if the merry-go-round started from rest? Ignore any friction in the axle about which the platform rotates. ••
58. Earth has an inertia of 5.97×10^{24} kg and a radius of 6.37×10^6 m. (*a*) Calculate its rotational inertia about the rotation axis, assuming uniform density. (*b*) The rotational inertia of Earth about its rotation axis has been measured to be 8.01×10^{37} kg · m^2. What does the difference between this value and the value you calculated in part *a* tell you about the distribution of inertia throughout the planet? ••
59. The rotational inertia of any flat object about an axis perpendicular to the object is equal to the sum of the rotational inertias about two mutually perpendicular axes in the plane of the object, if all three axes pass through a common point. Using this "perpendicular-axis theorem," determine (*a*) the rotational inertia of a uniform hoop of inertia m and radius R about a diameter of the hoop, and (*b*) the rotational inertia of a uniform square sheet of inertia m and side length a about a line drawn from the midpoint of one edge to the midpoint of the opposite edge. ••
60. The structure in Figure P11.60 is made of three identical rods of uniform density. Rank the rotational inertias of the structure, from smallest to largest, about the four rotation axes represented by the dashed lines. ••

Figure P11.60



61. Global warming might result in the melting of the polar ice caps. What effect would this melting have on the length of the day? ••

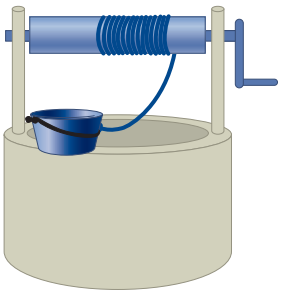
62. A hollow cylinder, a solid cylinder, and a billiard ball are all released at the top of a ramp and roll to the bottom without slipping. (*a*) Rank them, from least to greatest, according to the fraction of the kinetic energy that is rotational as they roll. (*b*) What are the ratios of their speeds when they reach the bottom of the ramp? ••

63. Estimate the rotational inertia of a 70-kg athlete about a horizontal axis passing through her waist from left to right. ••

64. A 0.20-kg block and a 0.25-kg block are connected to each other by a string draped over a pulley that is a solid disk of inertia 0.50 kg and radius 0.10 m. When released, the 0.25-kg block is 0.30 m off the ground. What speed does this block have when it hits the ground? ••

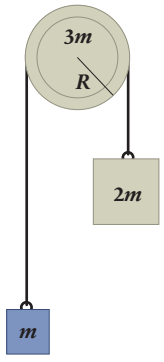
65. You accidentally knock a full bucket of water off the side of the well shown in Figure P11.65. The bucket plunges 15 m to the bottom of the well. Attached to the bucket is a light rope that is wrapped around the crank cylinder. How fast is the handle turning when the bucket hits bottom? The inertia of the bucket plus water is 12 kg. The cylinder has a radius of 0.080 m and inertia of 4.0 kg. ••

Figure P11.65



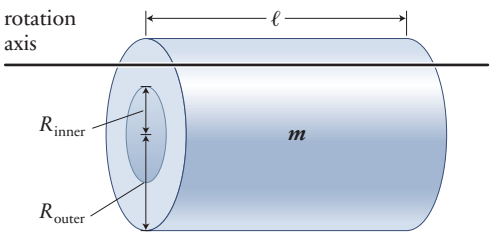
66. A block of inertia m is attached to a block of inertia $2m$ by a light string draped over a uniform disk of inertia $3m$ and radius R that can rotate on a horizontal axle (Figure P11.66). Initially, the block of inertia m is held in place so that the string is taut. When this block is released and so is free to move, what is its speed after it has risen a height h? Ignore friction between disk and axle. ••

Figure P11.66



67. A thick-walled cylinder of inertia m has an outer radius R_{outer}, an inner radius $R_{inner} = R_{outer}/2$, and a length ℓ (Figure P11.67). It rotates around an axis attached to its outside along its length. What is its rotational inertia about this axis? ••

Figure P11.67



68. A thin 0.15-kg rectangular sheet has 40-mm short sides and 80-mm long sides. Determine the rotational inertia of the sheet about an axis perpendicular to it and located (*a*) at the center of a long side, (*b*) at the center of a short side, and (*c*) at a corner. ••

69. One type of wagon wheel consists of a 3.0-kg hoop fitted with four 0.90-kg thin rods placed along diameters of the hoop so as to make eight evenly spaced spokes. For a hoop of radius 0.40 m, what is the rotational inertia of the wheel about an axis perpendicular to the plane of the wheel and through the center? ••

70. A thin rod, 0.72 m long, is pivoted such that it hangs vertically from one end. You want to hit the free end of the rod just hard enough to get the rod to swing all the way up and over the pivot. How fast do you have to make the end go? ••

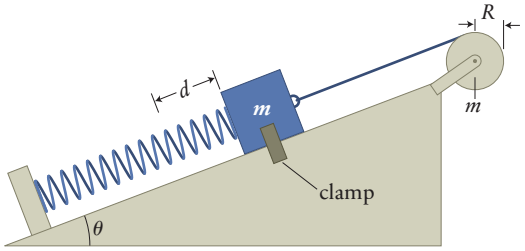
71. Big Ben in London is the most accurate mechanical clock of its size. The 300-kg hour hand is 2.7 m long, and the 100-kg minute hand is 4.2 m long. Calculate the rotational kinetic energy of the two-hand system, treating each hand as if it were a thin rod. ••

72. Consider an object that has rotational inertia I about some arbitrary axis of rotation. The *radius of gyration* of the object is the distance from this axis where all the object's inertia could be concentrated and still produce the same rotational inertia I about the axis. What is the radius of gyration of (*a*) a solid disk of radius R about an axis perpendicular to the plane of the disk and passing through its center, (*b*) a solid sphere of radius R about an axis through its center, and (*c*) a solid sphere of radius R about an axis tangent to the sphere? •••

73. A thin piece of metal in the shape of a right triangle has inertia m. Its perpendicular sides are of lengths ℓ and 2ℓ. Determine the rotational inertia about an axis that runs (*a*) along side ℓ and (*b*) along side 2ℓ. •••

74. A block of inertia m is attached to a light string wound around a uniform disk also of inertia m (Figure P11.74). The disk has radius R and rotates on a fixed horizontal axle through its center. The opposite side of the block is attached to a spring of force constant k and relaxed length ℓ attached at the bottom of an inclined surface that makes an angle θ with the horizontal. The block is at rest with the spring stretched a distance d from its relaxed length, with the block held in this position by a clamp. When the clamp is loosened so that the block is free to move, the block is pulled down the plane by the spring. What is the speed of the block when the spring is at half of its relaxed length? Ignore friction. •••

Figure P11.74



75. A certain baseball bat is a tapered cylinder that increases in diameter from 0.040 m at one end to 0.070 m at the other end and is 0.84 m long. The bat has uniform density and inertia 0.85 kg. What is the rotational inertia of this bat around an axis that passes through its skinny end and is perpendicular to the axis of the bat? •••

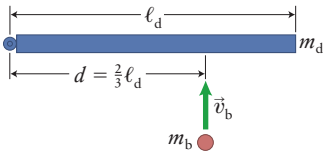
Additional Problems

76. In the laboratory, you are trying to put pennies on a rotating platform so that they don't slide off. Your partner tells you, "Pennies on the outside stay on better because $a_c = v^2/r$. For larger r, the centripetal acceleration is less, and so the frictional force has an easier time holding a penny in place out there." Is his reasoning correct? •
77. Solar and lunar eclipses are caused by a fortuitous relationship between the sizes of the Moon and Sun and their distances from Earth. Look up data for the Earth-Moon and Earth-Sun distances and the diameters of the Moon and the Sun. (*a*) Calculate the *angular size* of the Sun's disk as seen by an observer on Earth, where *angular size* is the angle *subtended* by the Sun's disk. (*b*) Calculate the angular size of the Moon's disk. •
78. Show that the centripetal acceleration of an object undergoing circular motion at constant speed can be expressed as

$$a_c = \frac{4\pi^2 r}{T^2},$$

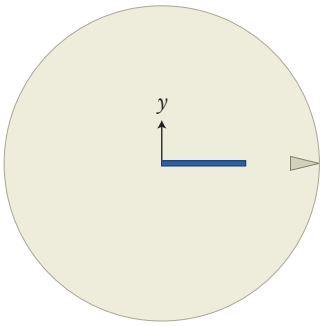
where r is the circle radius and T is the period of the motion. •
79. If you spin a raw egg, stop it momentarily, and then let go, it begins to spin again. You can't do this with a hard-boiled egg. Why? •
80. If you stretch your arm out full length and block the Sun with your thumb, how fast must your thumb move, in millimeters per minute, to track the Sun's motion? ••
81. Out looking for Tarzan, Jane swings on a vine to speed her search. (*a*) During her arc downward, how much work is done on her by the tensile force in the vine? (*b*) At which position along the arc is the force exerted by the vine on Jane greatest? ••
82. The Moon orbits Earth while Earth orbits the Sun, and thus the Moon's motion around the Sun is the combination of those two motions. Is the Moon's motion ever *retrograde* if observed from the Sun, where *retrograde* means that the Moon appears to move in the direction opposite the direction of Earth's motion relative to the distant stars? (You need to look up values for the Earth-Sun distance and the Earth-Moon distance to answer this question.) ••
83. Suppose several particles are in concentric circular orbits under the influence of some centripetal component of force, the magnitude of which depends on distance from the center of the motion. Each particle has a different orbital radius r and a corresponding period T. How does the force depend on r (*a*) if T is proportional to $\sqrt{r}$, (*b*) if T is proportional to $r^{3/2}$, and (*c*) if T is independent of r? ••
84. Approximate the Sun as a uniform sphere of radius 6.96×10^8 m, rotating about its central axis with a period of 25.4 days. If, at the end of its life, the Sun collapses inward to form a uniform dwarf star that is approximately the same size as Earth, what will the period of the dwarf's rotation be? ••
85. An open door of inertia m_d and width ℓ_d is at rest when it is struck by a thrown ball of clay of inertia m_b that is moving at speed v_b at the instant it strikes the door (Figure P11.85). The impact location is a distance $d = \frac{2}{3}\ell_d$ from the rotation axis through the hinges. The clay ball strikes perpendicular to the door face and sticks after it hits. What is the rotational speed ω of the door-ball system? ••

Figure P11.85



86. Your aunt owns an amusement park, and she wants you to add a circular loop to an existing roller-coaster ride. The first hill for the existing roller coaster is 55 m tall, and your aunt wants you to build, right after this hill, the tallest loop possible without having the cars fall out of the track or the passengers fall out of the cars. You think for a minute and realize what the minimum speed at the top of the loop has to be, and this gives you what you need to design the loop. ••• CR
87. The spacecraft in the movie *2001: A Space Odyssey* has a rotating cylinder to create the illusion of gravity, inside of which the crew walks and exercises. While watching this old movie, you realize that the radius of the cylinder is roughly three times a crew member's height. You imagine this cylinder rotating sufficiently rapidly to replicate Earth's gravity, but you worry about the difference between the acceleration at the top of the astronaut's head and the acceleration at her feet. Not wanting light-headed astronauts, you phone NASA with your new design. ••• CR
88. A turntable of radius 0.320 m is fitted with an electromagnet that causes steel objects to stay in place on the table. A triangular mark on the rim shows the rotational position of the table relative to the surroundings (Figure P11.88). The turntable completes one rotation in 0.360 s, and the rotation is counterclockwise in the view shown in the figure. A 0.125-kg steel rod of length 0.160 m is placed on the table so that one end is at the center and the rod is oriented radially outward, as shown in the figure. The electromagnet is switched on, and the turntable is started. After the table reaches its steady-state rotational speed, the electromagnet is switched off at an instant when the rod lies along the positive x axis. Describe the ensuing motion of the rod. •••

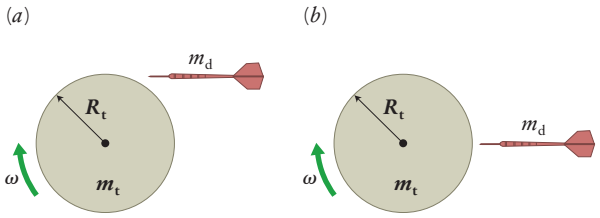
Figure P11.88



89. You have a toy top that plays music if the rotational speed is at least 1800 rev/min. The way to get this top spinning is to pull on a string wrapped around a spindle of diameter of 6.0 mm. The top has a rotational inertia of about 5.0×10^{-3} kg · m^2. The string is 1.2 m long. To hear the top sing, how hard must you pull on the string? ●●● CR

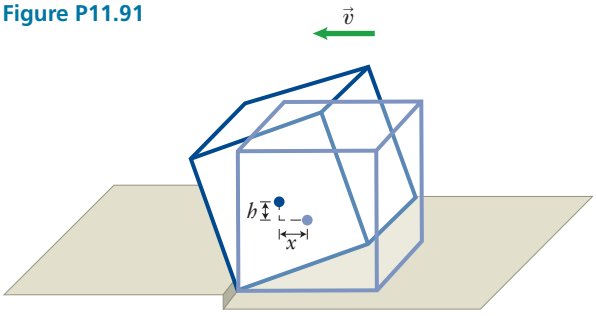
90. A dart of inertia m_d is fired such that it strikes with speed v_d, embedding its tip in the rim of a target that is a uniform disk of inertia m_t and radius R_t. The target is initially rotating clockwise in the view shown in Figure P11.90, with rotational speed ω about an axis that runs through its center and is perpendicular to its plane. Assume that the dart's inertia is concentrated at its tip. What is the final rotational speed of the target if the dart strikes (*a*) tangent to the target rim as in Figure P11.90*a* and (*b*) normal to the rim as in Figure P11.90*b*? ●●●

Figure P11.90



91. A smooth-faced cube of inertia m and side length d sliding along a highly polished floor at speed v strikes a door threshold, and the collision makes the cube tilt as shown in Figure P11.91. As the cube tips over, its center of mass is located at a variable horizontal distance x; $x = 0$ when the cube strikes the threshold. What is the rotational speed of the cube as a function of x? The rotational inertia of a cube about one edge is $\frac{2}{3}md^2$. [Hint: Call the vertical distance traveled by the center of mass h.] ●●●

Figure P11.91





12

Torque

CONCEPTS

12.1 Torque and angular momentum
12.2 Free rotation
12.3 Extended free-body diagrams
12.4 The vectorial nature of rotation

QUANTITATIVE TOOLS

12.5 Conservation of angular momentum
12.6 Rolling motion
12.7 Torque and energy
12.8 The vector product

Having described the basic principles of rotational motion, we now turn our attention to the cause of changes in this motion. Continuing to exploit the analogy between translational and rotational motion, we introduce a rotational analog of force called *torque*. A torque can cause an object's angular momentum to change, in the same way that a force can cause an object's momentum to change. We also study how to describe rotations that take place about an axis that is not fixed. To complete our description of rotations, we examine rotations in more than one dimension.

12.1 Torque and angular momentum

In the preceding chapter we studied how objects can be made to rotate. As I mentioned in Section 11.5, tangential components of forces cause the angular momentum of an object to change (Figure 11.31). If you exert a force on the edge of a stationary wheel in a direction tangential to the rim, as illustrated in **Figure 12.1a**, the wheel starts to rotate. To twist the cap off a jar, you exert a tangential force at the edge of the cap, as in Figure 12.1*b*.

When you exert a force on an object to set it in translational motion, only the magnitude and direction of that force play a role: The object accelerates in the direction of the force, and the greater your force, the greater the acceleration. When you exert a force on an object to twist or rotate the object, not only the magnitude and the direction of the force but also the point at which you exert it determine how effective that force is in rotating the object. For example, if you push on a seesaw to lift a child seated on the opposite end, as illustrated in **Figure 12.2**, it is best to push as far as possible from the pivot of the seesaw and to push in a direction perpendicular to the seesaw. If you push very close to the pivot or if you push in a direction not perpendicular to the seesaw, it is more difficult to rotate the seesaw about its pivot. For a force of a given magnitude, the ability to rotate the seesaw is greatest when the force is exerted as indicated in Figure 12.2*a*.

The reason that the ability of a force to rotate objects depends on the point of application is most easily understood in terms of energy and work. Suppose you raise the child of Figure 12.2 by exerting a force $\vec{F}$ straight down on the opposite end of the seesaw, as in **Figure 12.3*a*** on the next page. Raising the child a vertical distance Δy means increasing the gravitational potential energy of the Earth-child system by an amount $mg\,\Delta y$, where m is the child's inertia. This increase in potential energy is due to work done by you on the system. As you raise the child a distance Δy, your end of the seesaw goes down an equal amount. The work done by you on the system is equal to the magnitude of the force times the force displacement, or $F\,\Delta y_F$. If the child begins and ends at rest, so that there is no change in kinetic energy, the work done by you on the system must equal the change in potential energy of the Earth-child system: $mg\,\Delta y = F\,\Delta y_F$. So we see that to lift the child by pushing on the opposite end of the seesaw, the magnitude of the force $\vec{F}$ must be equal to the magnitude mg of the force of gravity on the child.

Suppose, however, you had pushed not on the far end of the seesaw but rather halfway between the pivot and the end. As Figure 12.3*b* shows, that point of the seesaw moves down only half the original force displacement, and so to do the same amount of work, the force you must exert is twice as great as before. The closer to the pivot you push, the greater the force required to raise the child.

We can use the same line of reasoning to understand why the force is most effective when it is oriented perpendicular to the seesaw. If the exerted force points in any other direction, we can decompose it into components

Figure 12.1 To (*a*) make a wheel rotate and (*b*) twist a cap off a jar, you must exert a force tangential to their rims.

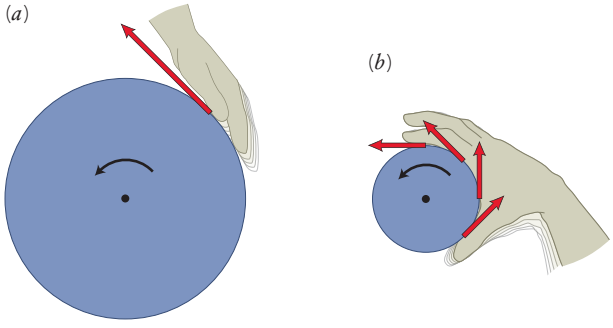
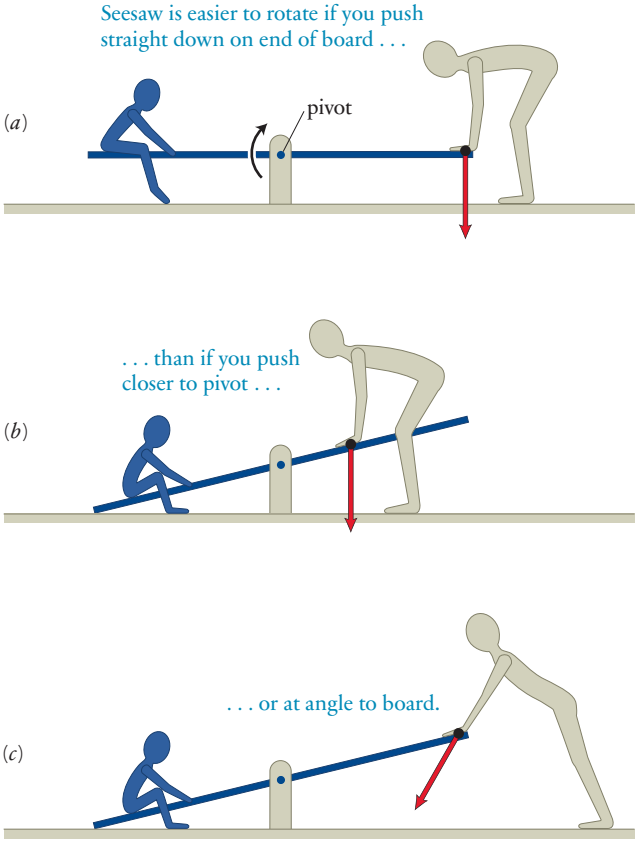
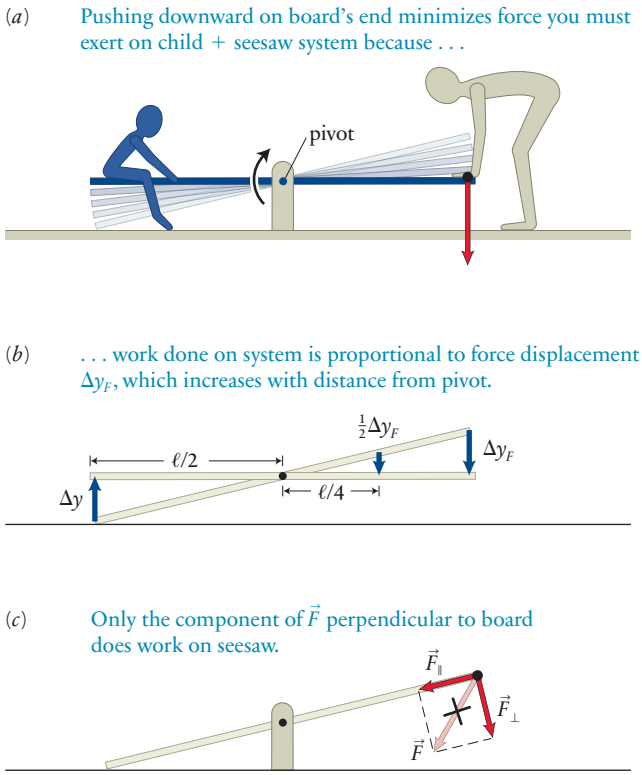


Figure 12.2 Easier and harder ways to rotate a seesaw.



CONCEPTS

Figure 12.3 The reason you can rotate the seesaw most easily by pushing downward on the end of the board.

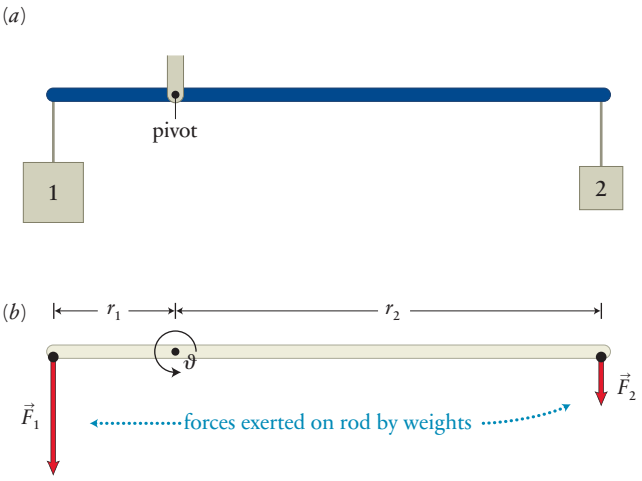


parallel and perpendicular to the seesaw (Figure 12.3*c*). The component parallel to the seesaw merely pushes the seesaw against the pivot and doesn't contribute to any rotational displacement of the seesaw: In the extreme case that the force is directed entirely along the seesaw, nothing happens. So it is only the component perpendicular to the seesaw, $\vec{F}_\perp$, that causes the seesaw to rotate.

To find out more accurately what determines the turning ability of a force, consider the rod in **Figure 12.4**. The rod is suspended by a pivot attached at a point that is off-center, and the rod is free to rotate about that point. Two objects of unequal inertias m_1 and m_2 are suspended on either end of the rod, keeping the rod in balance so that it is at rest. For simplicity we assume that the inertia of the rod is negligible compared to m_1 and m_2. Before reading on, complete the following checkpoint.

12.1 (*a*) Draw a free-body diagram for the rod of Figure 12.4. Let the inertia of the rod be negligible compared to m_1 and m_2. (*b*) Would the free-body diagram change if you slide object 2 to the left? (*c*) Experiments show that when $m_1 = 2m_2$ the rod is balanced for $r_2 = 2r_1$. How is the ratio r_1/r_2 related to the ratio m_1/m_2? (*d*) What would happen if you remove object 1 from the rod? (*e*) What would happen if you double the inertia of object 1? (*f*) What would happen if you carefully place a penny on top of object 1? Let the inertia of the two objects be significantly larger than that of the penny. (*g*) Is there a difference between what would happen in parts *e* and *f*?

Figure 12.4 Unequal weights balance a rod that is suspended off-center.



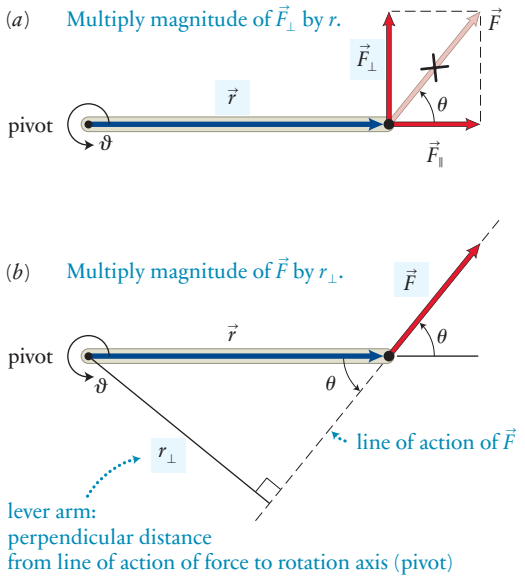
The forces exerted by the two objects in Figure 12.4 each have the ability to rotate the rod—object 1 tends to rotate the rod in a counterclockwise direction, object 2 in a clockwise direction. This ability of a force to rotate an object about an axis is called the **torque** about that axis. Torque is the rotational analog of force: Forces cause (translational) acceleration; torques cause rotational acceleration about an axis. The rod in Figure 12.4 is balanced because the torques caused by the two forces—their ability to rotate the rod—are equal in magnitude, but the rotations they tend to cause are in opposite directions, and so they sum to zero.

Note that the magnitudes of the two forces that cause these equal torques in opposite directions are not themselves equal: $F_1 = m_1 g$ and $F_2 = m_2 g$, with $m_2 \neq m_1$. From experiment we know, however, that the rod is balanced when $r_1 F_1 = r_2 F_2$. This suggests that the torque caused by each force is $r_\perp F$, the product of the magnitude of that force and the lever arm distance of the force relative to the axis of rotation (the perpendicular distance between the line of action of the force and the pivot; see Section 11.5). The line of action of the upward force exerted by the pivot on the rod (see Checkpoint 12.1) goes through the axis of rotation. Consequently, the corresponding lever arm distance is zero, and so this force does not cause any torque about the pivot.

In Figure 12.4, the forces exerted on the rod are perpendicular to it. **Figure 12.5** shows two equivalent ways to determine the torque caused by a force that is not perpendicular to the radial line from the pivot to the point of application of the force. One way is to decompose the force into components parallel and perpendicular to this axis (Figure 12.5*a*). Only the perpendicular component causes the lever to rotate; the parallel component simply pulls on the pivot. The torque caused by $\vec{F}$ is thus $rF_\perp = r(F \sin\theta)$, where r is the length of the vector pointing from the axis of rotation to the point of application of the force. Note that we can also write

CONCEPTS

Figure 12.5 Two ways to determine the torque caused by a force exerted on a lever at an angle θ.



this torque as $(r \sin\theta)F$, which is precisely $r_\perp F$, as we surmised. Figure 12.5*b* shows how to determine $r_\perp$: It is the perpendicular distance between the line of action of the force and the axis of rotation. So, the torque caused by a force exerted on an object is the product of the magnitude of the force and its lever arm distance. It can be written equivalently as $rF_\perp$ and as $r_\perp F$.

Like other rotational quantities, torque carries a sign that depends on the choice of direction for increasing ϑ. In Figure 12.4, for example, the torque caused by $\vec{F}_1$ about the pivot tends to rotate the rod in the direction of increasing ϑ and so is positive; the torque caused by $\vec{F}_2$ is negative. The sum of the two torques about the pivot is then $r_1F_1 + (-r_2F_2)$. As we've seen, the two torques are equal in magnitude when the rod is balanced, and so the sum of the torques is zero. When the sum of the torques is not zero, the rod's rotational acceleration is nonzero, and so its rotational velocity and angular momentum change.

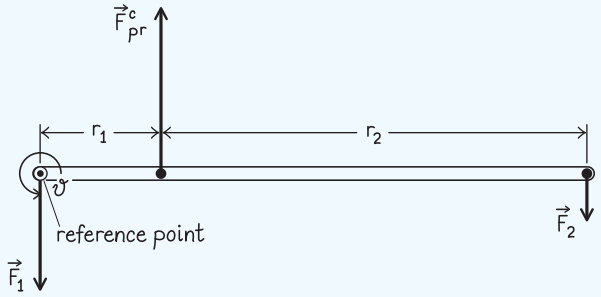
In the situations depicted in Figures 12.4 and 12.5 we used the pivot to calculate the lever arm distances. This is a natural choice because that is the point about which the object under consideration is free to rotate. However, torques also play a role for stationary objects that are suspended or supported at several different points and that are not free to rotate—for example, a plank or bridge supported at either end. To determine what reference point to use in such cases, complete the following exercise.

EXERCISE 12.1 Reference point

Consider again the rod in Figure 12.4. Calculate the sum of the torques about the left end of the rod.

SOLUTION I begin by making a sketch of the rod and the three forces exerted on it, showing their points of application on the rod (**Figure 12.6**).

Figure 12.6



The lever arm distances must now be determined relative to the left end of the rod. The lever arm distance of force $\vec{F}_1$ to this point is zero, and so the torque caused by that force about the left end of the rod is zero. If I choose counterclockwise as the positive direction of rotation, $\vec{F}_2$ causes a negative torque about the left end of the rod; the force $\vec{F}^c_{pr}$ exerted by the pivot causes a positive torque about the left end of the rod. The lever arm distance of $\vec{F}_2$ about the left end of the rod is $r_1 + r_2$; that of $\vec{F}^c_{pr}$ is r_1. Because the rod is at rest, the magnitude of the force exerted by the pivot is equal to the sum of the forces $\vec{F}_1$ and $\vec{F}_2$. Taking into account the signs of the torques, we find that the sum of the torques about the left end of the rod is $r_1(F_1 + F_2) - (r_1 + r_2)F_2 = r_1F_1 - r_2F_2$. This is the same result we obtained for the torques about the pivot, and so the sum of the torques about the left end is zero. ✔

Exercise 12.1 shows that the sum of the torques about the left end of the rod is zero, just like the sum of the torques about the pivot. You can repeat the calculation for the torques about the right end of the rod or any other point, and each time you will find that the sum of the torques is zero. The reason is that the rod is not rotating about any point, and so the sum of the torques must be zero about any point. In general we can say:

For a stationary object, the sum of the torques is zero.

For a stationary object we can choose any reference point we like to calculate torques. It pays to choose a reference point that simplifies the calculation. As you have seen, we do not need to consider any force that is exerted at the reference point. So, by putting the reference point at the point of application of a force, we can eliminate that force from the calculation.

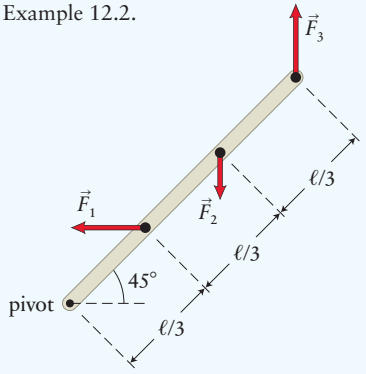
12.2 In the situation depicted in Figure 12.2*a*, you must continue to exert a force on the seesaw to keep the child off the ground. The force you exert causes a torque on the seesaw, and yet the seesaw's rotational acceleration is zero. How can this be if torques cause objects to accelerate rotationally?

CONCEPTS

EXAMPLE 12.2 Torques on lever

Three forces are exerted on the lever of Figure 12.7. Forces $\vec{F}_1$ and $\vec{F}_3$ are equal in magnitude, and the magnitude of $\vec{F}_2$ is half as great. Force $\vec{F}_1$ is horizontal, $\vec{F}_2$ and $\vec{F}_3$ are vertical, and the lever makes an angle of 45° with the horizontal. Do these forces cause the lever to rotate about the pivot? If so, in which direction?

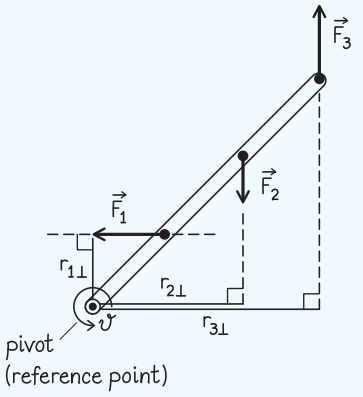
Figure 12.7 Example 12.2.



❶ **GETTING STARTED** I begin by arbitrarily choosing counterclockwise as the direction of increasing ϑ. With that choice of ϑ, $\vec{F}_1$ and $\vec{F}_3$ cause positive torques about the pivot, while $\vec{F}_2$ causes a negative torque. To answer the question, I need to determine the magnitude and sign of the sum of these three torques about the pivot.

❷ **DEVISE PLAN** The forces are not perpendicular to the long axis of the lever, and so I need to follow one of the two procedures shown in Figure 12.5 to determine the torques about the pivot. I arbitrarily choose to determine the lever arm distances. To determine these distances relative to the pivot, I make a sketch showing the forces and the perpendicular distance from the pivot to the line of action of each force (Figure 12.8). I can then get the magnitude of each torque by multiplying each force magnitude by the corresponding lever arm distance. Knowing the sign and magnitude of each torque, I can determine the combined effect of the three torques.

Figure 12.8



❸ **EXECUTE PLAN** I know that $F_1 = 2F_2$. My sketch tells me that $r_{1\perp} = (\ell/3)\sin 45°$ and $r_{2\perp} = (2\ell/3)\cos 45°$, and so $r_{1\perp} = \frac{1}{2}r_{2\perp}$ because $\sin 45° = \cos 45°$. Therefore the torques caused by these two forces about the pivot are equal in magnitude: $r_{1\perp}F_1 = \left(\frac{1}{2}r_{2\perp}\right)(2F_2) = r_{2\perp}F_2$. Because the two torques carry opposite signs, their sum is thus zero and their effects cancel. This means that the torque caused by $\vec{F}_3$ determines whether or not the lever rotates and, if so, in which direction. Because this torque is nonzero and counterclockwise, the lever rotates in a counterclockwise direction. ✔

❹ **EVALUATE RESULT** Looking at Figure 12.7, I see that the two larger forces ($\vec{F}_1$ and $\vec{F}_3$) cause counterclockwise torques about the pivot, and only the smaller force $\vec{F}_2$ causes a clockwise torque. Thus it makes sense that the lever rotates in the counterclockwise direction.



12.3 (*a*) Without changing the magnitude of any of the forces in Example 12.2, how must you adjust the direction of $\vec{F}_3$ to prevent the lever from rotating? (*b*) If, instead of adjusting the direction of $\vec{F}_3$, you adjust the magnitude of $\vec{F}_2$, by what factor must you change it?

12.2 Free rotation

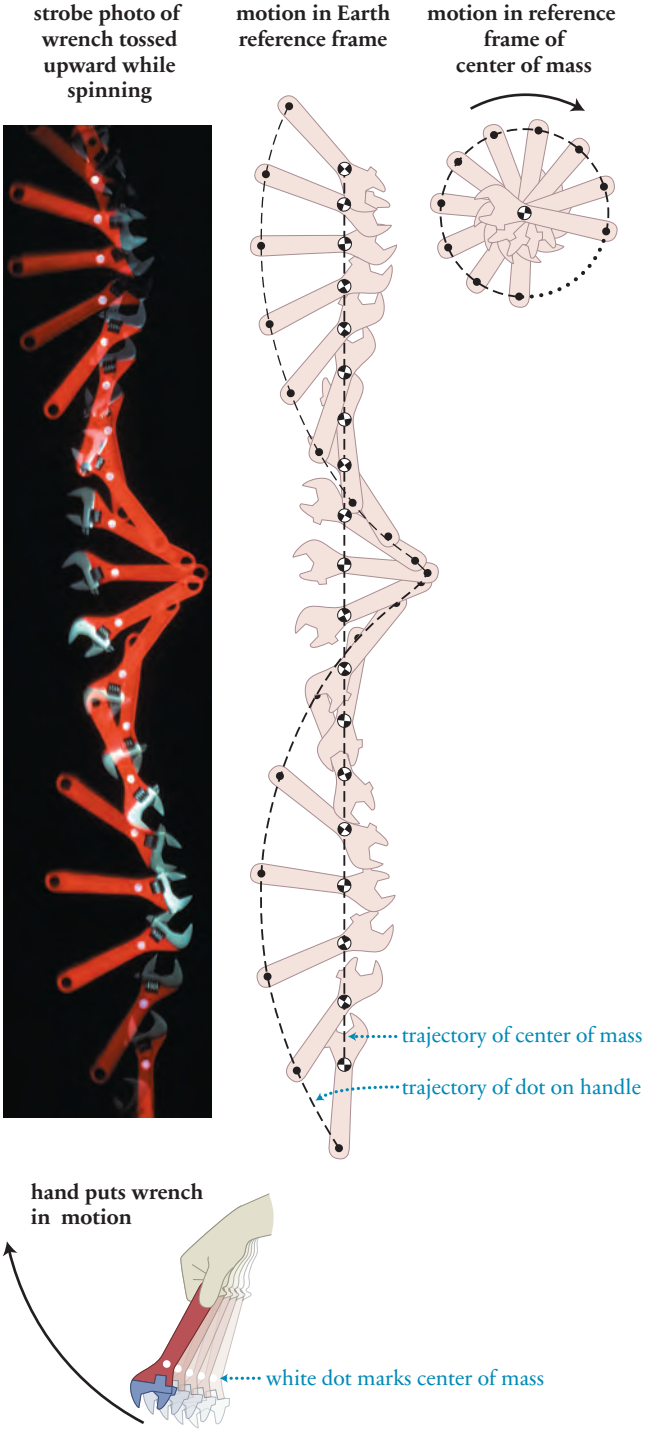
So far we have considered only rotations about a *fixed axis*, one that constrains the object to rotate about that axis and prevents it from moving away from the axis. Examples of such physical axes are axles, pivots, fulcrums, joints, and hinges. Objects don't need a physical axis to rotate, however. Put a pencil on your desk and flick one end with your finger. The pencil will move away from your hand while executing a rotating motion. A Frisbee flies through the air spinning about its center—without a physical axis through it. The rotation of objects that are not constrained by a physical axis or by other external constraints is called **free rotation.** About what point or axis do freely rotating objects like the pencil and the Frisbee rotate? Can an object be made to rotate about any point if nothing constrains it to rotate about some particular point?

Figure 12.9 shows the motion of a wrench after being launched as shown at the bottom of the figure. Hold the wrench by the end of the handle and sharply flip your wrist clockwise while launching the wrench upward. As the wrench moves upward, it executes a free rotation. The strobe photograph in Figure 12.9 allows us to analyze this motion. As you can see, the trajectory of the end of the handle is neither linear nor circular. Note, however, that the motion of the white dot on the wrench is remarkably simple: The dot moves straight up. What point does this white dot correspond to?

In Section 8.12, I showed that when external forces are exerted on the particles in a many-particle system, the system's center of mass moves as though all the particles in the system were concentrated at the center of mass and all external forces were exerted at that point. This statement is a direct result of conservation of momentum and of the definition of center of mass. The wrench, which consists of many atoms, is a many-particle system, and as it moves upward, Earth exerts a downward gravitational force on each atom in the wrench. The result we obtained in Section 8.12 thus states that the center of mass of the wrench should

CONCEPTS

Figure 12.9 Motion of a wrench that is launched upward so that it rotates.



move upward as if it were a particle moving under the influence of gravity. Indeed, if you closely examine the spacing of the dots in Figure 12.9, you can see that the spacing decreases. This tells you that the dot slows down as it moves upward. Quantitative analysis of the motion of the dot shows that the magnitude of the dot's downward acceleration is exactly g—that is, equal to that of an object launched without rotation.

So, the white dot on the wrench is the center of mass of the wrench, and other points on the wrench rotate around it. The center of mass is in translational motion at constant acceleration and thus cannot be rotating. This is true in general:

Objects that are made to rotate without external constraints always rotate about the center of mass.

This statement shows how remarkably powerful the concept of center of mass is. It allows us to break down the complex motion of the freely rotating wrench into two separate parts: the translational motion of the center of mass and a rotational motion of the wrench about the center of mass. This separation greatly simplifies the analysis of the motion because we can treat each part separately with the tools we have already developed.

12.4 As the wrench in Figure 12.9 moves upward, the upward translational motion of its center of mass slows down. Does the rotation about the center of mass also slow down? Which way does the wrench rotate when it falls back down after reaching its highest position?

12.3 Extended free-body diagrams

In Chapter 8 we developed a procedure for making free-body diagrams. These diagrams help us determine the *vector sum of the forces exerted on* an object, which, along with initial position and initial velocity, determines the motion of the object's center of mass. Free-body diagrams, however, do not provide any insight into the rotation of an object. For instance, if you push against the right end of a pencil that lies on a table, the pencil moves and rotates counterclockwise; pushing on the left end makes it rotate clockwise. If you push at the center of the pencil, you can even make it move without rotation. The free-body diagram for the pencil, however, is the same regardless of where you push.

To account for the rotation of an object, we must determine what torques are caused by the forces exerted on that object, and so we must know the lever arm distance of each force—information that is not contained in the free-body diagram because all forces are drawn with their tail at the center of the object. We must therefore develop a new kind of diagram that shows not only the forces exerted on an object but also the location of the point of application of each force relative to some chosen reference point. Such a diagram is called an **extended free-body diagram.** The procedure for making these diagrams is described in the Procedure box on page 394.

Drawing an extended free-body diagram and choosing a reference point help you determine the lever arm distances to that point so you can calculate the magnitude and direction of the torques about that point.

CONCEPTS

PROCEDURE: Extended free-body diagrams

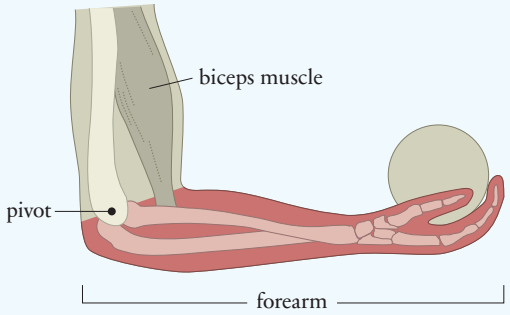
1. Begin by making a standard free-body diagram for the object of interest (the *system*) to determine what forces are exerted on it. Determine the direction of the acceleration of the center of mass of the object, and draw an arrow to represent this acceleration.
2. Draw a cross section of the object in the plane of rotation (that is, a plane perpendicular to the rotation axis) or, if the object is stationary, in the plane in which the forces of interest lie.
3. Choose a reference point. If the object is rotating about a hinge, pivot, or axle, choose that point. If the object is rotating freely, choose the center of mass. If the object is stationary, you can choose any reference point. Because forces exerted at the reference point cause no torque, it is most convenient to choose the point where the largest number of forces is exerted or where an unknown force is exerted. Mark the location of your reference point and choose a positive direction of rotation. Indicate the reference point in your diagram by the symbol ⊙.
4. Draw vectors to represent the forces that are exerted on the object and that lie in the plane of the drawing. Place the tail of each force vector at the point where the force is exerted on the object. Place the tail of the gravitational force exerted by Earth on the object at the object's center of mass.* Label each force.
5. Indicate the object's rotational acceleration in the diagram (for example, if the object accelerates in the positive ϑ direction, write $\alpha_\vartheta > 0$ near the rotation axis). If the rotational acceleration is zero, write $\alpha_\vartheta = 0$.

*Is the gravitational force really exerted at the center of mass? Suppose the gravitational force exerted by Earth were exerted at some other point. The force would then cause a permanent torque about the center of mass, and any object dropped from rest would begin spinning, which is not true.

EXERCISE 12.3 Holding a ball

You hold a ball in the palm of your hand, as shown in **Figure 12.10**. The bones in your forearm act like a horizontal lever pivoted at the elbow. The bones are held up by the biceps muscle, which makes an angle of about 15° with the vertical. Draw an extended free-body diagram for your forearm.

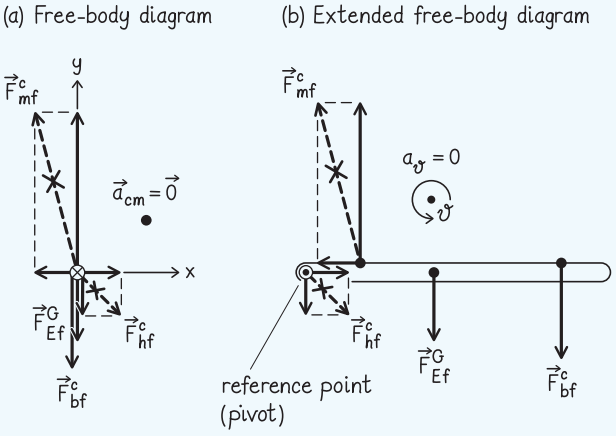
Figure 12.10 Exercise 12.3.



SOLUTION I begin by making a standard free-body diagram for the forearm, which is in contact with the upper-arm bone (the *humerus*), the biceps muscle, and the ball, all of which exert contact forces on it. In addition, there is the gravitational force $\vec{F}^G_{Ef}$ to show in my diagram. I know the direction of $\vec{F}^c_{bf}$ (down), $\vec{F}^c_{mf}$ (up and to the left), and $\vec{F}^G_{Ef}$ (down) but not the direction of $\vec{F}^c_{hf}$.

How do I come up with a reasonable direction for $\vec{F}^c_{hf}$ in my diagram? I know that the forearm is at rest, which means its linear acceleration is zero, and so the vector sum of the forces exerted on the forearm must be zero. This means that the component of $\vec{F}^c_{hf}$ parallel to the forearm must be equal in magnitude to the parallel component of $\vec{F}^c_{mf}$ but pointing in the opposite direction. Because I cannot determine off-hand what the perpendicular component of $\vec{F}^c_{hf}$ is, I draw a force vector pointing down at some arbitrary angle and adjust its length so that the two parallel components are equal in length. Finally I adjust all the perpendicular components so that their sum is zero. This gives me the free-body diagram shown in **Figure 12.11a**. Because the arm is at rest, I write $\vec{a}_{cm} = \vec{0}$ next to the diagram. The zero acceleration tells me that the vector sum of the forces exerted on the arm is zero.

Figure 12.11



Next I turn to the extended free-body diagram. I represent the forearm as a horizontal rod, with the pivot on the left, as shown in Figure 12.11*b*. Because the forearm is at rest, I can choose any point as the reference point. Because I know the least about $\vec{F}^c_{hf}$, however, it makes sense to choose my reference point at the pivot where that force is exerted. I choose the positive direction of rotation to be counterclockwise and write $\alpha_\vartheta = 0$ next to it because the forearm is not rotating.

Finally I draw the same forces as in the free-body diagram, placing their tails at their points of application. The result is shown in Figure 12.11*b*. The perpendicular component of $\vec{F}^c_{mf}$ causes a positive torque, $\vec{F}^G_{Ef}$ and $\vec{F}^c_{bf}$ cause negative torques, and $\vec{F}^c_{hf}$ causes no torque because its lever arm distance is zero. ✔

You will work out a quantitative solution to this problem in Checkpoint 12.8.

CONCEPTS

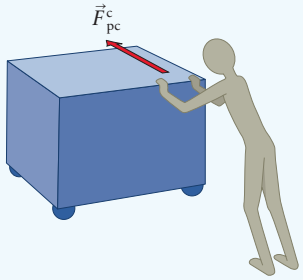
12.5 (*a*) If the biceps muscle in Figure 12.10 were attached farther out toward the wrist, would the torque generated by the muscle about the pivot get greater, get smaller, or stay the same? (*b*) As the hand is raised above the level of the elbow, so that the forearm makes an angle of 15° with the horizontal, does the arm's capacity to lift objects increase, decrease, or stay the same?

As you may have noticed, the information in the extended free-body diagram overlaps with that in the free-body diagram. Is it then necessary to draw both? I strongly recommend doing so until you have gained some practice. The standard free-body diagram helps you find the vector sum of the forces exerted on the object, which determines the object's translational motion. In the extended free-body diagram, it is more cumbersome to decompose vectors and add components, and so it is easier to make errors. The extended free-body diagram, however, is necessary to determine the torques caused by the forces that are exerted on the object. The next exercise helps clarify the purpose of each diagram.

EXERCISE 12.4 Pushing a large crate

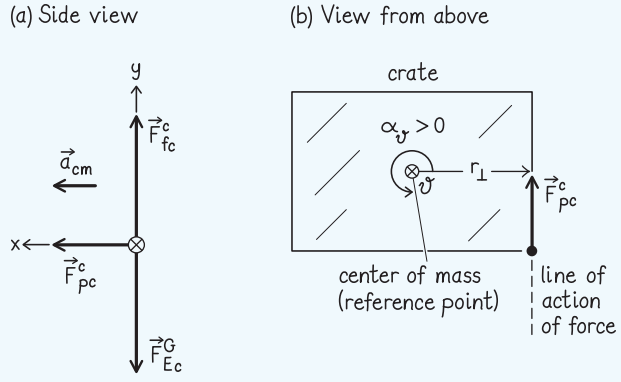
You are moving a large crate mounted on swivel wheels, exerting an off-center force $\vec{F}^{c}_{pc}$ as shown in **Figure 12.12**. Draw an extended free-body diagram for the crate. You can ignore any friction in the wheels.

Figure 12.12 Exercise 12.4.



SOLUTION The forces exerted on the crate are a downward force of gravity, an upward contact force exerted by the floor, and your horizontal push. I therefore draw the free-body diagram shown in **Figure 12.13*a***.

Figure 12.13



The free-body diagram shows that the vector sum of the forces exerted on the crate is nonzero. From this information I conclude that the push causes the center of mass of the crate to accelerate in the direction of the push. I therefore draw an arrow next to the free-body diagram to indicate the direction of this center-of-mass acceleration.

The crate is not constrained to rotate about a fixed axis. The floor and the wheels it is resting on, however, do constrain it to rotate in the horizontal plane. In addition to accelerating the center of mass of the crate, the off-center push therefore causes the crate to rotate about a vertical axis through the center of mass. Because the rotation is in the horizontal plane, I draw an outline of the crate as seen from above for the extended free-body diagram.

The crate is not constrained to rotate about any fixed axis in the plane I have drawn, so the center of mass is the reference point I must use. I assume the crate to be uniform, so its center of mass is at the center. I mark the reference point in my diagram and choose counterclockwise as the direction of positive rotation.

The only force that lies in this plane is $\vec{F}^{c}_{pc}$, so the extended free-body diagram contains just one force (Figure 12.11*b*). The lever arm distance of $\vec{F}^{c}_{pc}$ is the distance between its line of action and the center of mass of the crate. From the diagram, I thus conclude that $\vec{F}^{c}_{pc}$ causes a counterclockwise torque. This torque causes the crate to rotate counterclockwise, and I write $a_\vartheta > 0$ next to the diagram to indicate that the rotational acceleration of the crate is positive. ✔

12.6 Suppose the force $\vec{F}^{c}_{pc}$ in Exercise 12.4 gives the center of mass of the crate an acceleration $\vec{a}_{cm}$. Would this acceleration be greater, smaller, or the same if the force were exerted exactly at the center of the crate?

12.4 The vectorial nature of rotation

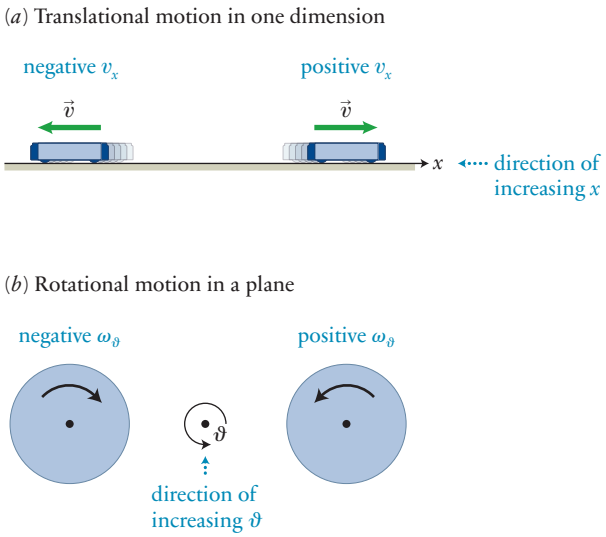
So far we've treated all rotational quantities as if they were scalars. There are two reasons a scalar treatment of rotational quantities cannot be complete, however: Rotations have direction, and the rotation axis has a definite orientation in space. For rotations that lie in a single plane, an algebraic sign is sufficient to indicate the direction of rotation in the same way that an algebraic sign indicates direction in one-dimensional translational motion, as shown in **Figure 12.14** on the next page.

For rotations in three dimensions, the algebraic sign alone is not sufficient to determine the direction of rotation. Consider, for example, the two spinning disks in **Figure 12.15** on the next page. If you view them from the angle shown in Figure 12.15*a*, you could claim that the disk A spins counterclockwise and disk B clockwise. Seen from below, as in Figure 12.15*b*, however, disk A spins clockwise and the rotation of disk B, now seen sideways, is neither clockwise nor counterclockwise. So, in three dimensions the terms *clockwise* and *counterclockwise* are insufficient to specify the direction of a rotation, just as *positive* and *negative* are not enough to specify a displacement vector in more than one dimension.

Can a rotation be characterized by a vector? The vectors $\vec{r}$, $\vec{v}$, and $\vec{a}$ of an object in circular motion at

CONCEPTS

Figure 12.14 For both translation in one dimension and rotation in a plane, the two possible directions of motion can be distinguished by their algebraic sign.



constant speed all lie in the plane of the trajectory, as shown in **Figure 12.16*a***. Their directions change every instant, and because all directions in the plane are equivalent, it makes no sense to associate a vector in this plane with the rotation. How would we agree on which direction to pick? The direction of the rotation axis, a direction in which the vectors $\vec{r}$, $\vec{v}$, and $\vec{a}$ never point, is the only direction associated with the rotation that remains fixed. We can specify the orientation of the rotation axis by a vector. Given that vector, you know the orientation of the plane in which the rotation takes place.

If we use a vector to specify the orientation of the rotation axis in space, we can use the two possible directions along this axis to represent the two possible directions of rotation about the axis. The convention for associating a vector along the rotation axis with the direction of rotation is illustrated in Figure 12.16*b* and is called the **right-hand rule:**

> **When you curl the fingers of your right hand along the direction of rotation, your thumb points in the direction of the vector that specifies the direction of rotation.**

Figure 12.15 For rotations in three dimensions, it is no longer possible to specify the direction of rotation unambiguously by algebraic sign or by the terms *clockwise* and *counterclockwise*.

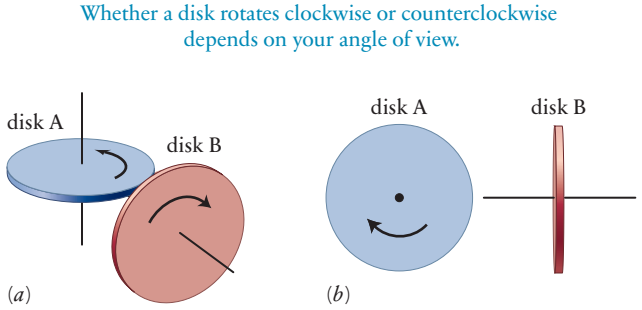
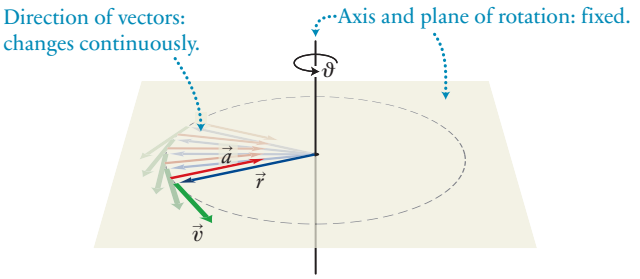


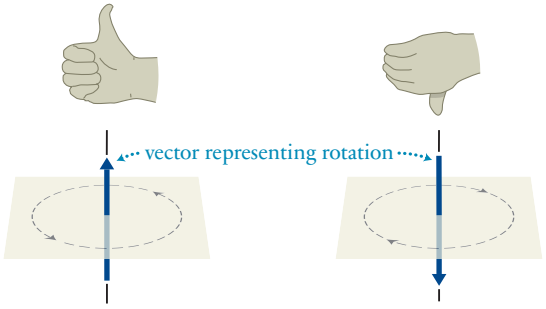
Figure 12.16 Right-hand rule and directions associated with rotations.

(*a*) Vectors $\vec{r}$, $\vec{v}$, and $\vec{a}$ cannot be associated with a direction of rotation.



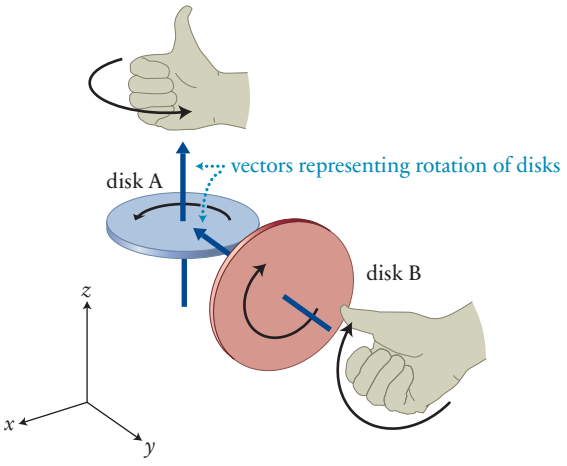
(*b*) The right-hand rule connects direction of rotation about an axis with a direction along the rotation axis.

Right-hand rule: Curl fingers of right hand in direction of rotation; thumb will point in direction of vector representing rotation.



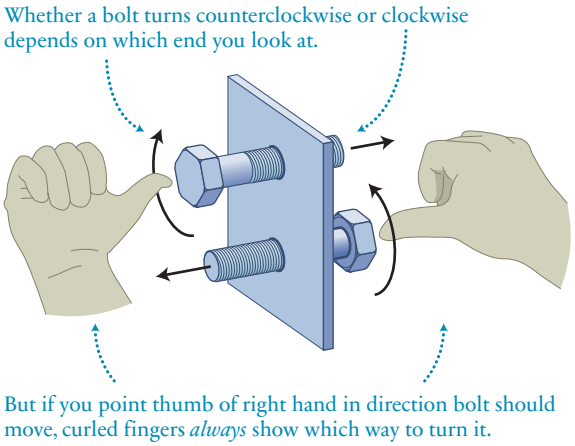
Armed with this rule, you can now specify the direction of any rotation and the orientation of the axis about which that rotation takes place with a vector and avoid the ambiguity of terms like *clockwise* and *counterclockwise*. For example, for the two disks in Figure 12.15*a* the right-hand rule yields the directions given by the two blue vectors in **Figure 12.17**. The advantage of the vectors is that we can specify them by giving their x, y, and z components. The vector representing

Figure 12.17 The right-hand rule applied to the spinning disks of Figure 12.15.



CONCEPTS

Figure 12.18 The right-hand rule tells you how to tighten a bolt regardless of your position relative to the bolt.



the rotation of disk A points in the positive z direction, and so it has the form $(0, 0, +A)$. The vector representing the rotation of disk B has the form $(0, -B, 0)$, where A and B are positive numbers.

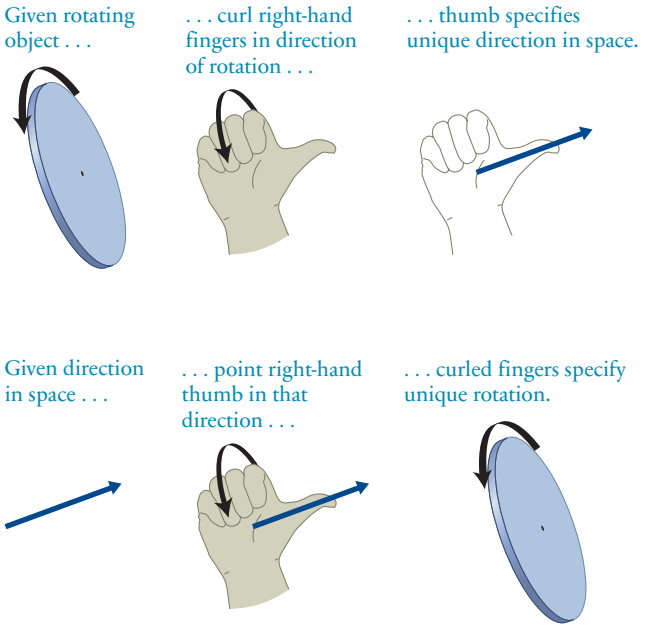
Note that the vector associated with a rotation always lies along the axis of that rotation. If you invert the direction of rotation, the vector points in the opposite direction along the axis. For example, if disk A in Figure 12.17 were rotating in the opposite direction from the one illustrated, you'd have to curl your fingers in the opposite direction, and so your thumb would point in the negative z direction, opposite the direction shown.

Regardless of your observation angle, the right-hand rule always yields the same direction for a given rotation. Suppose, for example, that you are tightening bolts into a plate as in **Figure 12.18**. When you tighten a bolt, it is natural to consider the bolt's rotation as being clockwise because you usually look at the head of the bolt when tightening it. Now imagine tightening a bolt on the backside of the plate by reaching around it. From where you stand, you must turn in a *counterclockwise* direction to tighten that bolt! The statement "to tighten a bolt, you must rotate it in a clockwise direction" holds only when you are facing the head of the bolt. As Figure 12.18 shows, however, when you align your right-hand thumb in the direction you want the bolt to move, your fingers curl in the correct direction, regardless of where you stand relative to the plate.

The right-hand rule can be used to determine not only the direction of a vector representing a rotation but also the other way around. Given a vector, we can use the right-hand rule to determine the corresponding rotation (**Figure 12.19**). Either one uniquely determines the other.

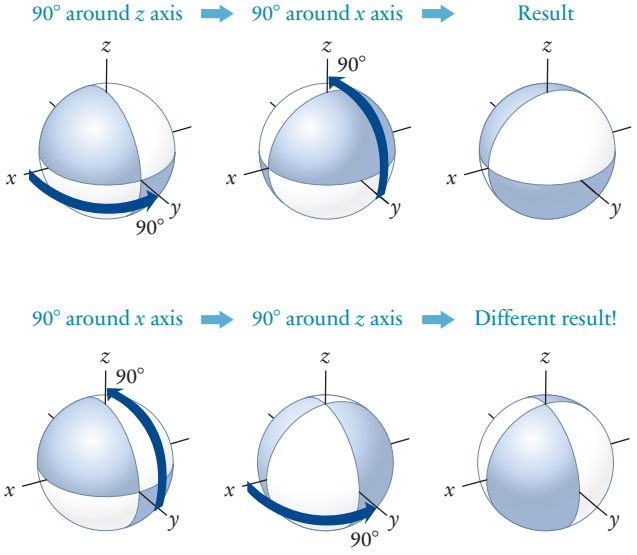
What kind of a vector is the vector we have just introduced? All the vector quantities we have encountered up to now are derived from the displacement vector—an arrow that points from an object's initial

Figure 12.19 The right-hand rule used two ways.



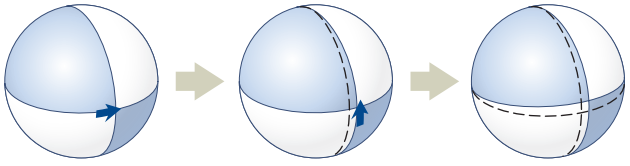
position to its final position. As we've seen, displacement vectors obey a commutation principle: The sum of several displacements is independent of the sequence in which the displacements are added. Displacement vectors commute in two and in three dimensions. Do rotational displacements commute? Rotations in a plane do, as you can easily verify by adding two angles. The result is independent of the order in which the two angles are added: $15° + 23° = 23° + 15° = 38°$. As **Figure 12.20** illustrates, however, rotations in three dimensions do not commute: If you give a beach ball two successive rotations over 90° angles about perpendicular axes, the final orientation of the ball depends on

Figure 12.20 Rotational displacements are not commutative: The order in which you perform them matters.

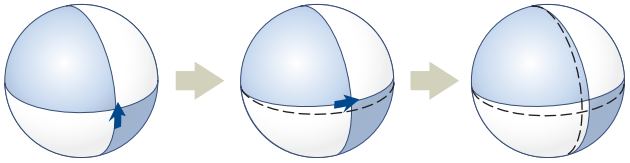


CONCEPTS

Figure 12.21 Successive rotations over very small angles are commutative. The dashed lines show the initial positions of the ball's seams.



For successive rotations over very small angles, order doesn't matter.



the sequence of the rotations. So, rotational displacements cannot be vectors because in three dimensions $\vartheta_1 + \vartheta_2 \neq \vartheta_2 + \vartheta_1$.

If, instead of rotating the ball twice over 90° angles, you rotate it over two small angles θ_1 and θ_2, the difference in the ball's final orientation is much smaller. As **Figure 12.21** shows, for two successive rotations over 6° angles, the difference is no longer noticeable. Why small angles commute can be understood by looking at **Figure 12.22**, which shows a close-up of the displacement of the seams on the ball in Figure 12.21. Each rotational displacement $\Delta\vartheta$ corresponds to a displacement over an arc of length $\Delta s = r\,\Delta\vartheta$, where r is the radius of the ball. For very small rotational displacements $d\vartheta_1$ and $d\vartheta_2$, the corresponding arcs ds_1 and ds_2 are approximately straight lines on a nearly flat surface. The smaller the angles θ_1 and θ_2, the more the arcs behave like vectors in two dimensions and so commute.

So, infinitesimally small rotational displacements do commute: $d\vartheta_1 + d\vartheta_2 = d\vartheta_2 + d\vartheta_1$ This means that rotational velocity—the rate at which the rotational displacement changes, $\omega = d\vartheta/dt$—obeys the commutation law: $\omega_1 + \omega_2 = \omega_2 + \omega_1$. We can therefore introduce a *rotational velocity vector* $\vec{\omega}$ whose direction is given by the right-hand rule and whose magnitude is ω. The rotational velocity vector is the basic vector describing rotational motion, much as displacement is the basic vector for translational motion. **Figure 12.23** shows two spinning tops with their rotational velocity vectors. Each vector determines the orientation of the top's axis of rotation, the direction in which the top rotates around this axis, and the rotational speed. The arrow representing the rotational velocity of top A is longer than the arrow of top B, indicating that A spins faster than B. We shall examine the vector nature of other rotational quantities in Section 12.8.

Vectors associated with a direction of rotation are called **axial vectors;** vectors derived from the displacement vector are called **polar vectors.** Axial vectors play an important role in physics, not only in rotation but also in electricity and magnetism.

12.7 Suppose the rotation of top A in Figure 12.23 slows down without a change in the direction of its axis of rotation. (*a*) In which direction does the vector $\Delta\vec{\omega}$ point? (*b*) Can the top's rotational acceleration be represented by a vector? If so, in which direction does this vector point?

CONCEPTS

Figure 12.22 Detail of the balls of Figure 12.21 in their final orientation. Because the displacements along the arcs corresponding to each rotation are nearly straight lines, the rotations commute.

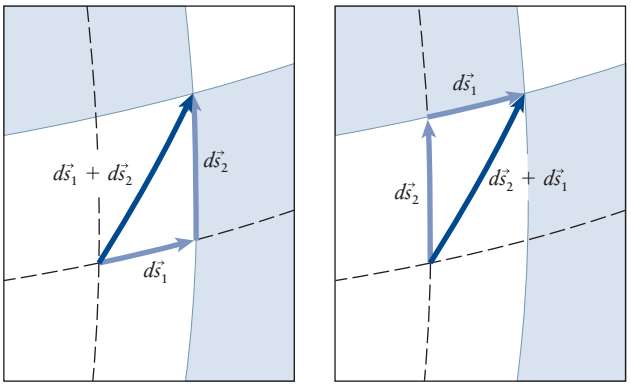
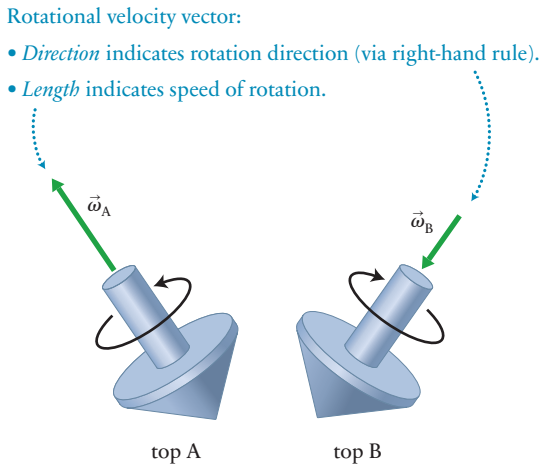


Figure 12.23 Representing rotational velocity vectors.



Self-quiz

1. A rope supports one end of a beam as shown in **Figure 12.24**. Draw the lever arm distance for the torque caused by the rope about the pivot.
2. Draw a free-body diagram and an extended free-body diagram for (*a*) a door hanging on two hinges and (*b*) a bridge supported from each end, with a car positioned at one-quarter of the bridge's length from one support.
3. Which diagram in **Figure 12.25**—1, 2, or 3—shows the alarm clock on the left after it has been rotated in the directions indicated by (*a*) 90° about the *x* axis and then 90° about the *y* axis and (*b*) 90° about the *y* axis and then 90° about the *x* axis? Does the order of the rotation change your answer?

Figure 12.24

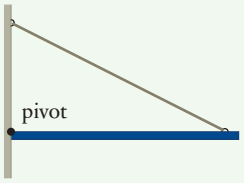
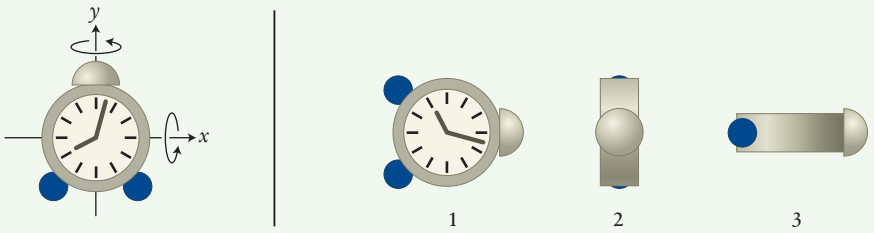
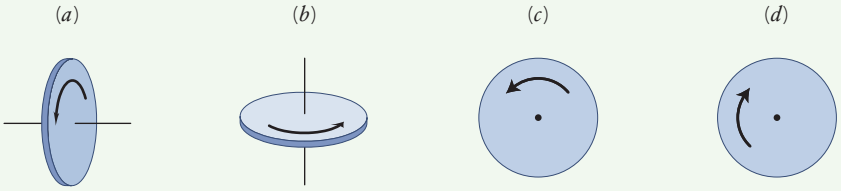


Figure 12.25



4. Give the direction of the rotational velocity vector associated with each spinning object shown in **Figure 12.26**.

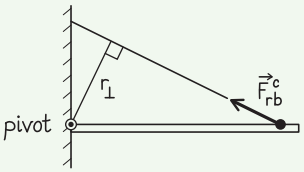
Figure 12.26



Answers

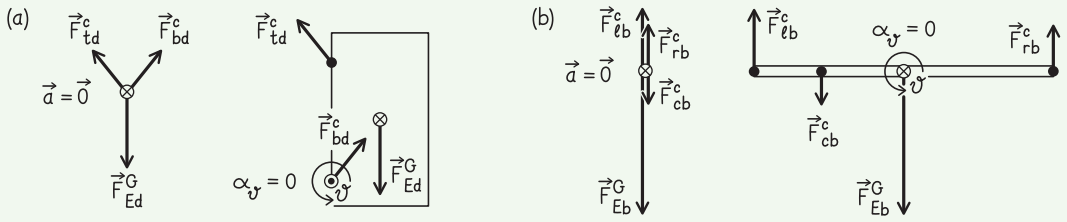
1. The lever arm distance $r_\perp$ is the perpendicular distance between the pivot and the line of action of the force exerted by the rope on the beam, as shown in **Figure 12.27**.

Figure 12.27



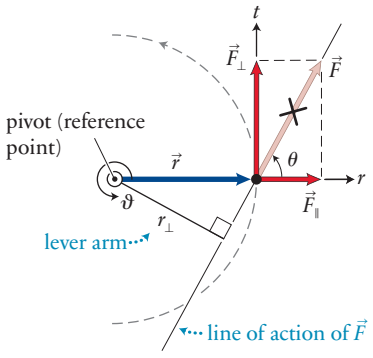
2. See **Figure 12.28**. (*a*) The door interacts with three objects: Earth, the top hinge, and the bottom hinge. Without the top hinge, the force of gravity would tend to rotate the door about an axis perpendicular to the door through the bottom hinge. The force exerted by the top hinge must balance the clockwise torque caused by the force of gravity about the axis through the bottom hinge. The horizontal components of the forces exerted by the hinges must cancel each other. (*b*) The bridge interacts with four objects: Earth, the right support, the left support, and the car. The upward forces from the supports must balance the downward gravitational forces of the car and the bridge. Because these forces must also counteract the counterclockwise torque caused by the car, the force exerted by the support closer to the car must be greater than the force exerted by the other support.

Figure 12.28



3. (*a*) 3; (*b*) 2. The order of rotation does make a difference.
4. Wrapping the fingers of your right hand in the direction of spin gives rotational velocity vectors that point (*a*) to the right, (*b*) up, (*c*) out of the page, and (*d*) into the page.

Figure 12.29 A force $\vec{F}$ is exerted on a particle that is constrained to move in a circle.



12.5 Conservation of angular momentum

In Section 11.5 we saw that angular momentum is a conserved quantity and that the angular momentum of a system rotating in a plane on which no tangential forces are exerted remains constant. Let us now examine how tangential forces are related to the concept of torque introduced in the first part of this chapter and derive mathematical expressions that relate the various rotational quantities and torque.

Consider the situation shown in **Figure 12.29**. A force $\vec{F}$ is exerted on a particle of inertia m that is constrained to move along a circular trajectory of radius r about a pivot. Because $\vec{F}$ has a component that is tangential to the circular trajectory, the particle speeds up. In the language we developed in the first part of this chapter, the particle's rotational acceleration is nonzero because $\vec{F}$ causes a torque on the particle about the pivot. The magnitude of this torque, which we denote by the Greek letter τ (*tau*), is given by

$$\tau \equiv rF \sin\theta = r_\perp F = rF_\perp, \tag{12.1}$$

where θ is the angle between the radius vector $\vec{r}$ and $\vec{F}$ and where $F_\perp$ is the magnitude of the tangential component of the force: $F_\perp = |F_t|$. Because torque is the product of a force and a distance, it has SI units of newton-meters (N · m). Work, being the product of displacement and force, has the same units. Torque and work represent two very different and unrelated quantities, however. Work corresponds to the transfer of energy by a force; torque is the ability of a force to change the rotational motion of an object. To avoid confusion, we shall always use joules for work and newton-meters for torque.

Let us now relate this torque to the particle's tangential acceleration a_t (which is perpendicular to the radius vector $\vec{r}$). This acceleration is given by

$$F_t = ma_t. \tag{12.2}$$

Combining Eqs. 12.1 and 12.2, we obtain

$$\tau_\vartheta = r(ma_t), \tag{12.3}$$

where the subscript ϑ means that τ_ϑ is a signed quantity. Although strictly speaking τ_ϑ is the component of a torque *vector,* we shall refer to it simply as torque. The algebraic sign of the torque is determined by the sign of the rotational acceleration it causes. The torque caused by the force $\vec{F}$ in Figure 12.29, for example, is positive because it causes the particle to accelerate in the positive ϑ direction ($a_t > 0$).

As we saw in Chapter 11, the tangential acceleration is related to the rotational acceleration by $a_t = r\alpha_\vartheta$, so we can write

$$\tau_\vartheta = rm(r\alpha_\vartheta) = mr^2\alpha_\vartheta. \tag{12.4}$$

The quantity mr^2 is the rotational inertia of the particle, so Eq. 12.4 becomes

$$\tau_\vartheta = I\alpha_\vartheta. \tag{12.5}$$

Equation 12.5 shows how the torque caused by a force exerted on a particle relates to the particle's rotational acceleration. Notice the remarkable parallel between Eqs. 12.2 and 12.5: Instead of "force equals inertia times acceleration," we have "torque equals rotational inertia times rotational acceleration." Torque is to rotation what force is to translation, and therefore Eq. 12.5 is sometimes called the *rotational equation of motion.*

To find a similar expression for extended objects, consider the rotating extended object illustrated in **Figure 12.30**. Imagine breaking down the object into small particles of inertia δm_n, with each particle located a distance r_n from the axis and having velocity $\vec{v}_n$. Let each particle be subject to a torque $\tau_{n\vartheta}$ about the axis; this torque is caused by forces exerted by surrounding particles and by external forces. For each particle we can write an expression analogous to Eq. 12.4:

$$\tau_{n\vartheta} = \delta m_n \, r_n^2 \, \alpha_{n\vartheta}. \tag{12.6}$$

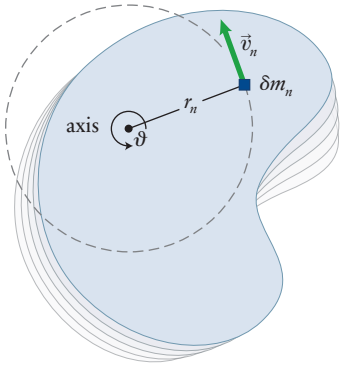
The sum of the torques on all particles is then

$$\sum_n \tau_{n\vartheta} = \sum_n (\delta m_n r_n^2 \alpha_{n\vartheta}) = \left(\sum_n \delta m_n r_n^2 \right) \alpha_\vartheta, \tag{12.7}$$

where we have used the fact that all particles in a rigid rotating object must have the same rotational acceleration α_ϑ. The sum on the right side is the rotational inertia I of the extended object (Eq. 11.42), and so Eq. 12.7 becomes

$$\sum_n \tau_{n\vartheta} = I\alpha_\vartheta. \tag{12.8}$$

Figure 12.30 A rigid object constrained to rotate about an axis.



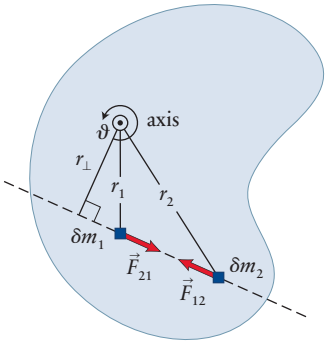
The sum on the left side contains torques caused by external forces (*external torques* for short) and internal torques. To find the contribution from internal torques, consider two interacting particles within the extended object, as illustrated in **Figure 12.31**. Because particles 1 and 2 interact with each other, they exert forces on each other that are equal in magnitude: $F_{12} = F_{21}$. As Figure 12.31 shows, the lever arm distance for the two forces is the same ($r_\perp$), and so the torque caused by $\vec{F}_{12}$ is equal in magnitude to the torque caused by $\vec{F}_{21}$ but opposite in direction. Their sum is therefore zero:

$$\tau_{21\vartheta} + \tau_{12\vartheta} = r_\perp F_{21} - r_\perp F_{12} = r_\perp (F_{21} - F_{12}) = 0. \tag{12.9}$$

Because this cancellation holds for all pairs of interacting particles within the object, all internal torques drop out of the sum on the left side of Eq. 12.8. Only torques caused by external forces remain. The expression we obtain for an extended object is therefore identical to that for a particle:

$$\Sigma\tau_{\text{ext}\,\vartheta} = I\alpha_\vartheta. \tag{12.10}$$

Figure 12.31 The internal torques caused by two interacting particles within a rotating rigid object have the same lever arm distance, so they cancel.



Another useful expression can be obtained by recognizing that the particle's rotational velocity and rotational acceleration are related by

$$\alpha_\vartheta = \frac{d\omega_\vartheta}{dt}. \tag{12.11}$$

Substituting this expression into Eq. 12.10 and recalling that angular momentum $L_\vartheta = I\omega_\vartheta$ (Eq. 11.34), we get

$$\Sigma\tau_{\text{ext}\,\vartheta} = I\frac{d\omega_\vartheta}{dt} = \frac{d}{dt}(I\omega_\vartheta) = \frac{dL_\vartheta}{dt}, \tag{12.12}$$

where we have used the fact that the rotational inertia I of a rigid object is constant. So torque is the time rate of change of angular momentum, much

as force is the time rate of change of momentum. Equation 12.12 shows that if the sum of the torques caused by external forces exerted on an object is zero, the object's angular momentum does not change:

$$\Sigma\tau_{\text{ext}\,\vartheta} = \frac{dL_\vartheta}{dt} = 0 \quad \Rightarrow \quad \Delta L_\vartheta = 0. \tag{12.13}$$

Whenever the sum of the torques caused by the forces exerted on an object is zero—in other words, whenever its angular momentum is constant—the object is said to be in **rotational equilibrium.** In Chapter 8 we learned that an object is in *translational equilibrium* when its momentum is constant and that this condition is met when the vector sum of the external forces exerted on the object is zero. Now we can expand that notion to include angular momentum, which remains constant when the sum of the torques caused by external forces exerted on it is zero. An object that is in both translational and rotational equilibrium is said to be in **mechanical equilibrium:**

$$\Sigma\tau_{\text{ext}\,\vartheta} = 0 \quad \text{and} \quad \Sigma\vec{F}_{\text{ext}} = \vec{0} \quad \Leftrightarrow \quad \text{mechanical equilibrium.} \tag{12.14}$$

For a system that is not in rotational equilibrium, we have

$$\Delta L_\vartheta = J_\vartheta, \tag{12.15}$$

where J_ϑ represents the transfer of angular momentum from the environment to the system. The quantity J_ϑ is called the **rotational impulse** delivered to the system. Like angular momentum it has SI units of $\text{kg}\cdot\text{m}^2/\text{s}$. In analogy to the momentum law (Eq. 4.18) and the energy law (Eq. 9.1), which embody conservation of momentum and energy, respectively, Eq. 12.15 embodies conservation of angular momentum and is called the **angular momentum law.**

To calculate the rotational impulse, we can use Eq. 12.12. For a system that is subject to constant torques, the time rate of change of the angular momentum is constant, so over a time interval Δt,

$$\Sigma\tau_{\text{ext}\,\vartheta} = \frac{\Delta L_\vartheta}{\Delta t}$$

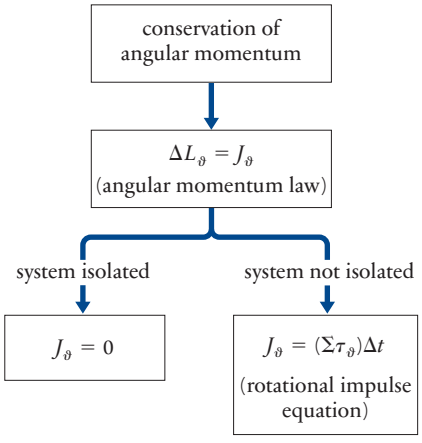
$$\Delta L_\vartheta = \left(\Sigma\tau_{\text{ext}\,\vartheta}\right)\Delta t. \tag{12.16}$$

Comparing Eqs. 12.15 and 12.16, we see that the rotational impulse is

$$J_\vartheta = \left(\Sigma\tau_{\text{ext}\,\vartheta}\right)\Delta t \quad \text{(constant torques)}. \tag{12.17}$$

Equation 12.17, called the **rotational impulse equation,** allows us to calculate the change in angular momentum for a system subject to constant external torques. **Figure 12.32** illustrates how conservation of angular momentum gives rise to the angular momentum law and how to treat isolated and nonisolated systems. Notice that the structure of this figure is identical to that of Figure 9.16, which illustrates the same principles for momentum and energy.

Figure 12.32 Angular momentum change for isolated and nonisolated systems.



12.8 Use the conditions of mechanical equilibrium to express $\vec{F}^{\text{c}}_{\text{hf}\,y}$ in Figure 12.11*b* in terms of $\vec{F}^{G}_{\text{Ef}}$ and $\vec{F}^{\text{c}}_{\text{bf}}$. Let the distance between the ball and the pivot be ℓ and the distance from the point where the biceps muscle is attached to the pivot be $\ell/5$.

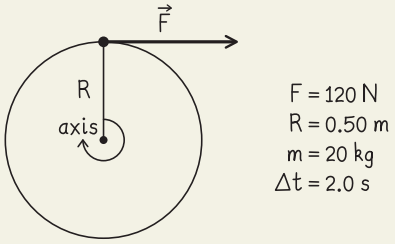
QUANTITATIVE TOOLS

EXAMPLE 12.5 Flywheel

A motor exerts a constant force of 120 N tangential to the rim of a 20-kg cylindrical flywheel of radius 0.50 m. The flywheel is free to rotate about an axis through its center and runs perpendicular to its face. If the flywheel is initially at rest and the motor is turned on for 2.0 s, how much work does the motor do on the flywheel?

❶ **GETTING STARTED** I begin by making a sketch of the situation to organize the information (**Figure 12.33**). The force exerted by the motor causes the flywheel to start spinning, which means the wheel's rotational kinetic energy changes. This is the only energy change in the system, and I know from the energy law (Eq. 9.1) that $\Delta E = W$. Therefore $\Delta K_{\text{rot}} = W$, and so to calculate the work done by the motor, I need to determine this change in rotational kinetic energy. Because the flywheel is at rest initially, I know that $\Delta K_{\text{rot}} = K_{\text{rot,f}}$.

Figure 12.33



❷ **DEVISE PLAN** To obtain $K_{\text{rot,f}}$, I can use Eq. 11.31, $K_{\text{rot}} = \frac{1}{2}I\omega^2$. The rotational inertia I of the flywheel (which is a solid cylinder) is $\frac{1}{2}mR^2$ (see Table 11.3). Because I'm interested in $K_{\text{rot,f}}$, I need the final value for ω, the wheel's rotational speed. How can I connect ω_{f} to anything I know in this problem? The relationship between ω and angular momentum is $L_\vartheta \equiv I\omega_\vartheta$ (Eq. 11.34), and I know from Eq. 12.16 that $\Delta L_\vartheta = \tau_\vartheta \Delta t$. I know Δt, but do I know anything about τ in terms of the information given—a force, an inertia, and a wheel radius? Yes, Eq. 12.1: $\tau = rF_\perp$. Thus my plan is to express ω_{f} in terms of R and F and then use that expression for ω_{f} in $K_{\text{rot}} = \frac{1}{2}I\omega^2$ to calculate $K_{\text{rot,f}} = W$.

❸ **EXECUTE PLAN** The magnitude of the torque caused by the motor is $\tau = RF$, where R is the radius of the wheel and F is the magnitude of the force exerted by the motor. Equation 12.16 then gives

$$\Delta L_\vartheta = \left(\Sigma\tau_{\text{ext}\,\vartheta}\right)\Delta t = +RF\Delta t.$$

Because the initial angular momentum is zero, I know that $\Delta L_\vartheta = I\omega_{\vartheta,\text{f}} = +I\omega_{\text{f}}$, and so I have

$$\Delta L_\vartheta = RF\Delta t = I\omega_{\text{f}}$$

$$\omega_{\text{f}} = \frac{RF\Delta t}{I}.$$

The final rotational kinetic energy is thus

$$K_{\text{rot,f}} = \tfrac{1}{2}I\omega_{\text{f}}^2 = \tfrac{1}{2}I\left(\frac{RF\Delta t}{I}\right)^2$$

$$= \frac{(RF\Delta t)^2}{2I} = \frac{(RF\Delta t)^2}{mR^2} = \frac{(F\Delta t)^2}{m},$$

and the work done on the flywheel is

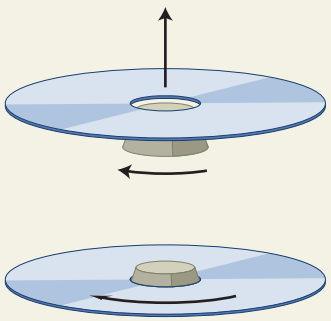
$$W = K_{\text{rot,f}} = \frac{[(120\text{ N})(2.0\text{ s})]^2}{20\text{ kg}} = 2880\text{ J} = 2.9\text{ kJ}. \checkmark$$

❹ **EVALUATE RESULT** Delivering 2.9 kJ in 2.0 s corresponds to a power of (2.9 kJ)/(2.0 s)=1.4 kW, which is not an unreasonable amount for a large motor.

EXAMPLE 12.6 Spinning disc

Consider the spinning disc shown in **Figure 12.34**, in which, a spinning conical shaft rises up into the opening in the center of the disc, and the disc begins to spin. Suppose the disc's rotational inertia is I_{d}, that of the shaft is I_{s}, and the shaft's initial rotational speed is ω_{i}. Does the rotational kinetic energy of the disc-shaft system remain constant in this process? Assume for simplicity that no external forces cause torques on the shaft.*

Figure 12.34 Example 12.6.



❶ **GETTING STARTED** I am given rotational inertias for a disc and a shaft, plus the initial rotational speed of the shaft, and my task is to determine whether or not the system's rotational kinetic energy changes when these two units interact: $\Delta K_{\text{rot}} \stackrel{?}{=} 0$. The disc is initially at rest, but as the shaft comes in contact with it, the two exert on each other forces that cause torques. The shaft speeds up the rotation of the disc, and the disc slows down the rotation of the shaft. The disc has no rotational velocity before the shaft touches it, and so initially all the system's rotational kinetic energy is in the shaft. After they reach a common rotational speed ω_{f}, both have rotational kinetic energy. I need to calculate the initial and final rotational kinetic energies of the shaft-disc system.

*In reality, the shaft would have to be connected to a motor that keeps it rotating at the correct rotational speed, and thus there would be a force that causes a torque on the shaft.

(Continued)

2 DEVISE PLAN The rotational kinetic energy of a rotating object is given by Eq. 11.31, $K = \frac{1}{2}I\omega^2$. I know the initial rotational speed of the shaft and its rotational inertia, so I can use this equation to calculate the initial rotational kinetic energy of the shaft-disc system. I do not know the final rotational speed of the system, but I do know that because there are no external torques on the system, Eq. 12.13 tells me that the angular momentum must remain constant: $\Delta L_\vartheta = 0$. Expressing ΔL_ϑ as the difference between the final and initial values gives me an expression containing ω_f and ω_i, which means I can probably get an expression for ω_f/ω_i that I can then use to compare the ratio $K_{rot,f}/K_{rot,i}$ and thereby determine whether or not $\Delta K_{rot} = 0$. Because the problem is stated in symbols rather than numerical values, my comparison will be between two algebraic expressions.

3 EXECUTE PLAN Because the torques that the disc and shaft cause on each other are internal and because there are no external torques, I have for the system's angular momentum

$$\Delta L_\vartheta = (I_s + I_d)\omega_{\vartheta,f} - I_s\,\omega_{\vartheta,i} = 0. \qquad (1)$$

If I let the initial direction of rotation of the shaft be positive, $\omega_{\vartheta,i} = +\omega_i$ and so $\omega_{\vartheta,f}$ is also positive. Rearranging terms in Eq. 1, I find that the ratio of the final and initial rotational speeds is

$$\frac{\omega_f}{\omega_i} = \frac{I_s}{I_s + I_d}.$$

The system's initial rotational kinetic energy is $K_i = \frac{1}{2}I_s\,\omega_i^2$, its final rotational kinetic energy is $K_f = \frac{1}{2}(I_s + I_d)\omega_f^2$, and the ratio of the two is

$$\frac{K_f}{K_i} = \frac{\frac{1}{2}(I_s + I_d)}{\frac{1}{2}I_s}\frac{\omega_f^2}{\omega_i^2} = \frac{I_s}{I_s + I_d} < 1,$$

so $K_f < K_i$, or $\Delta K \neq 0$. The rotational kinetic energy of the system is not constant. ✔

4 EVALUATE RESULT The spinning up of the disc is like an inelastic "rotational collision": The disc initially at rest comes in contact with the spinning shaft, and the two reach a common rotational speed. While the disc is spinning up, some of the system's initial rotational kinetic energy is converted to thermal energy because of friction between disc and shaft, and so it makes sense that the system's rotational kinetic energy decreases.

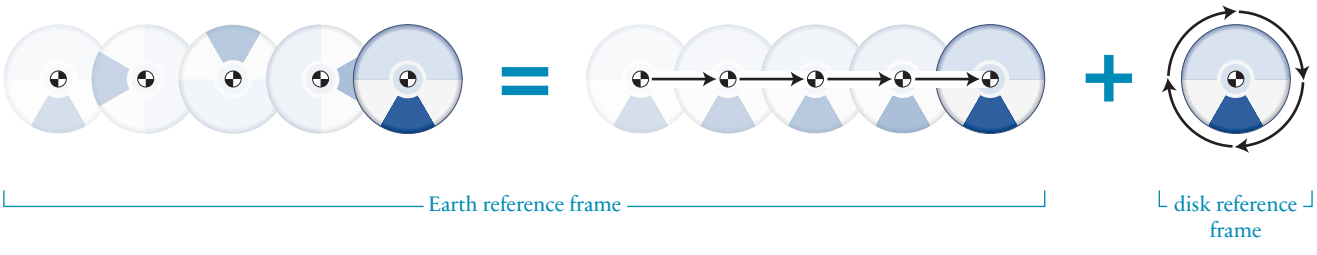
12.9 Consider the situation in Example 12.6. (*a*) Is the vector sum of the forces exerted by the shaft on the disc nonzero while the disc is spinning up? (*b*) Is the disc isolated?

12.6 Rolling motion

The rotations we have considered so far fall into two classes: rotations about a fixed axis and free rotations. Forces exerted on objects that are constrained to rotate about a fixed axis cause rotational motion only because the axis prevents any translational motion. Forces exerted on freely rotating objects cause both translational motion of the center of mass and rotation about that point. These two motions are *independent* of each other. When you throw a Frisbee, for example, its translational speed v_{cm} is independent of its rotational speed ω (**Figure 12.35**). The Frisbee can have any combination of spin and translational speed: You can throw it so that it spins fast and moves away from you slowly, spins slowly and moves away from you fast, and so on.

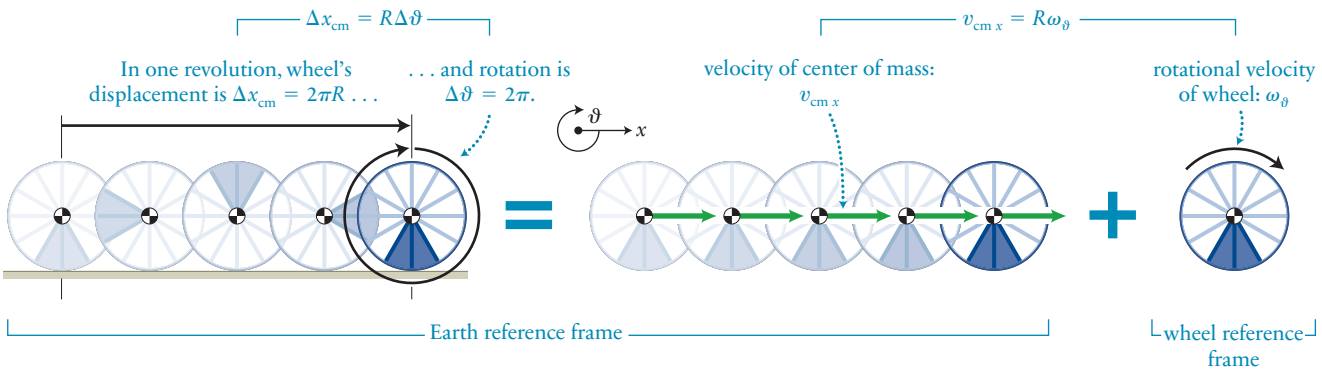
The rolling motion of a wheel or ball represents an intermediate case between fixed and free rotations because rolling puts a constraint on the relationship between the translational and rotational motions. Furthermore, a rolling wheel or ball always rotates about its geometric center. We shall restrict our discussion to symmetric objects for which the center of mass is at the geometric center.

Figure 12.35 The translational and rotational motions of a thrown Frisbee are independent of each other.



QUANTITATIVE TOOLS

Figure 12.36 When an object rolls without slipping, its translational and rotational motions are coupled.



Consider the wheel of radius R rolling along a straight line on a flat surface in **Figure 12.36**. To describe this rolling motion we need to choose both a direction of rotation and a direction of the x axis. Usually it is advisable to choose these directions such that a positive change in ϑ corresponds to a positive change in x, as shown in Figure 12.36. If the wheel does not slip, its center advances a distance equal to its circumference, $2\pi R$, during each full revolution. Because the wheel's rotational displacement in one full revolution is 2π, the center-of-mass displacement Δx_{cm} caused by the rolling is related to the rotational displacement $\Delta\vartheta$ by

$$\Delta x_{cm} = R\Delta\vartheta. \tag{12.18}$$

Note that this expression is valid for any displacement; for example, for a quarter revolution $\Delta\vartheta = 2\pi/4 = \pi/2$ and $\Delta x_{cm} = 2\pi R/4 = (\pi/2)R = R\Delta\vartheta$. Dividing the left and right sides by the time interval Δt it takes to advance this distance and taking the limit of both sides for $\Delta t \to 0$ yield

$$v_{cm\,x} = R\omega_\vartheta \quad \text{(rolling motion without slipping)}. \tag{12.19}$$

This equation expresses the **condition for rolling motion without slipping.** By coupling the translational and rotational speeds, this condition puts a constraint on the motion of any object that rolls without slipping. Unlike a Frisbee, whose (translational) speed is independent of its rotational speed, the speed of a moving car that is not skidding is proportional to the rotational speed of its wheels.

Equation 12.19 resembles Eq. 11.10 relating the translational and rotational speeds of an object in circular motion:

$$v_t = r\omega_\vartheta. \tag{12.20}$$

These two equations mean very different things, however. Equation 12.19 describes the translational motion of the wheel in the Earth reference frame, whereas Eq. 12.20 gives the tangential speed of any point on the rotating wheel in a reference frame moving along with the center of mass of the wheel. The velocity of a point on the wheel in the Earth reference frame is obtained by adding these two velocities, as shown in **Figure 12.37**. A point on the rim

Figure 12.37 The trajectory of a point on the rim of a rolling wheel. Such a trajectory is called a *cycloid*.

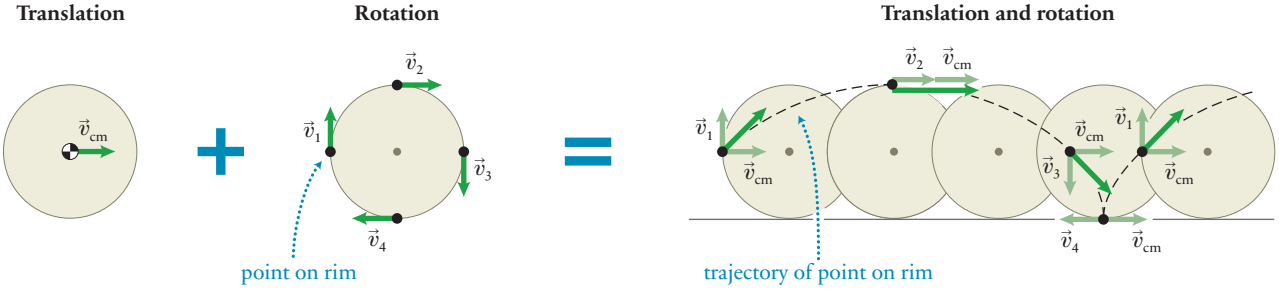
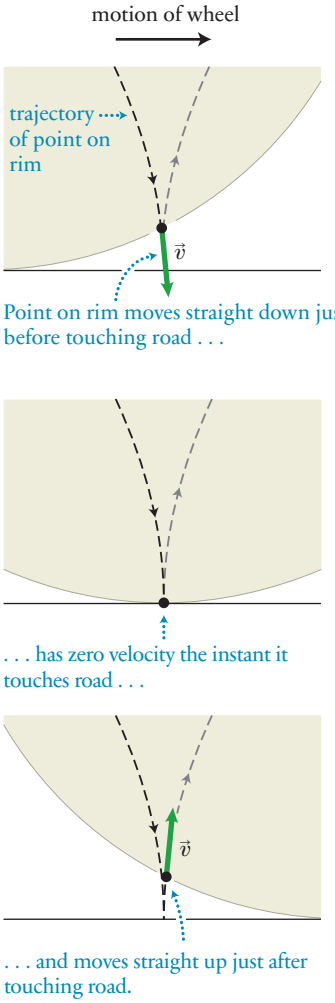


Figure 12.38 The point of contact of a wheel rolling without slipping has zero instantaneous velocity.



moves with a continuously changing velocity along a trajectory resembling that of the handle of the wrench in Figure 12.9. Note in particular that whenever a point on the rim is in contact with the surface over which the wheel is rolling, that point has zero instantaneous velocity (on the rim $r = R$ and $v_t = r\omega_\vartheta$ points in the direction opposite the direction of $v_{cm} = R\omega$). This zero instantaneous velocity of the rim point in contact with the surface is a direct consequence of the requirement that the object rolls without slipping: The relative velocity of the two surfaces in contact must be zero for there to be no slipping. **Figure 12.38** shows that points on the rim move in a direction perpendicular to the surface just before and just after reaching the bottom.

Let us next turn to the dynamics of rolling motion. Suppose a round object of inertia m and radius R is released from rest on a ramp as shown in **Figure 12.39a**. Why does the object roll down instead of sliding down as would a block of wood? The reason is that the force of static friction, which forces the instantaneous point of contact to be motionless, causes a torque about the center of the object. To understand the motion of the object, we must find the vector sum of the forces exerted on it and the sum of the torques caused by these forces. Figure 12.39*b* shows a free-body diagram for the object, which is subject to two forces: the force of gravity and a contact force exerted by the ramp, which has a normal component $\vec{F}^n_{ro}$ in the y direction and a tangential component $\vec{F}^s_{ro}$ in the x direction due to static friction. As we have done with other problems involving ramps, we chose the x axis to be along the ramp. For the object to accelerate down the ramp, the vector sum of the forces must point in the positive x direction, so $\Sigma F_y = F^G_{Eoy} + F^n_{ro} = 0$ and $\Sigma F_x = F^G_{Eox} - F^s_{ro} > 0$. The vector sum of the forces and the center-of-mass acceleration are related by

$$\Sigma F_x = F^G_{Eox} - F^s_{ro} = mg \sin\theta - F^s_{ro} = ma_{cmx}. \qquad (12.21)$$

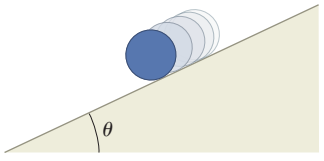
This equation contains two unknowns: the magnitude of the force of static friction F^s_{ro} and a_{cmx}. All we know about F^s_{ro} is that it must be less than or equal to its maximum value $(F^s_{ro})_{max}$, but this fact doesn't help us solve Eq. 12.21. Let us therefore turn to the extended free-body diagram shown in Figure 12.39*c*, which shows the object with the forces exerted on it at their points of application. Because the object rotates about its center of mass, we determine the sum of the torques about the center of mass. Neither $\vec{F}^G_{Eo}$ nor $\vec{F}^n_{ro}$ causes a torque because their lines of action go through the center of mass, and so the lever arm distances for these forces are zero. Only $\vec{F}^s_{ro}$ causes a torque; its lever arm distance is the radius R of the object. Equation 12.10 then yields

$$\Sigma\tau_{ext\,\vartheta} = +F^s_{ro}R = I\alpha_\vartheta, \qquad (12.22)$$

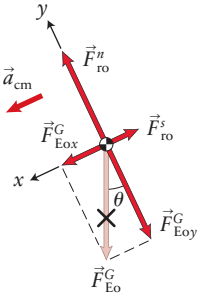
where I is the object's rotational inertia and α_ϑ is its rotational acceleration. Although this yields another equation, we have now added a third unknown: α_ϑ.

Figure 12.39 A round object rolls down a ramp.

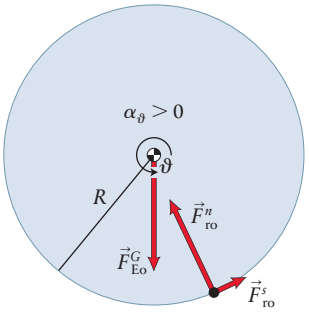
(*a*) Object rolls down ramp



(*b*) Free-body diagram



(*c*) Extended free-body diagram



QUANTITATIVE TOOLS

Because the object is rolling without slipping, its rotational and translational motions are coupled, and so its rotational and center-of-mass accelerations must be related. Differentiating both sides of Eq. 12.19, $v_{\mathrm{cm}x} = R\omega_\vartheta$, with respect to time yields

$$a_{\mathrm{cm}x} = R\alpha_\vartheta. \tag{12.23}$$

Substituting α_ϑ from Eq. 12.23 into Eq. 12.22 and solving for F^s_{ro}, we find

$$F^s_{\mathrm{ro}} = \frac{I}{R^2} a_{\mathrm{cm}x}. \tag{12.24}$$

Substituting this result for F^s_{ro} into Eq. 12.21 and solving for $a_{\mathrm{cm}x}$, we obtain

$$a_{\mathrm{cm}x} = +\frac{g\sin\theta}{1+\dfrac{I}{mR^2}} = +\frac{g\sin\theta}{1+c}, \tag{12.25}$$

where $c \equiv I/mR^2$ is the *shape factor* of the object ($0 < c \leq 1$; see Table 11.3). The numerator in Eq. 12.25 is the acceleration of an object sliding down a ramp with negligible friction (see Eq. 3.20). The denominator shows by what factor this acceleration is reduced for rolling motion. Because c is independent of m and R, Eq. 12.25 shows that the acceleration of the object does not depend on the values of m and R. Only the shape of the object matters because it affects the shape factor c. For all objects of a given shape, the time interval to roll down a ramp is thus the same. A large cylinder takes the same time interval as a small one, but a thin cylindrical shell takes longer than a solid cylinder because the shape factor of the shell is greater.

We can now also find an expression for the magnitude of the force of static friction $\vec{F}^s_{\mathrm{ro}}$. Substituting $a_{\mathrm{cm}x}$ from Eq. 12.25 into Eq. 12.24 and eliminating the rotational inertia using $c \equiv I/mR^2$, which gives $I = cmR^2$, we obtain

$$F^s_{\mathrm{ro}} = \frac{I}{R^2} a_{\mathrm{cm}x} = \frac{cmR^2}{R^2}\frac{g\sin\theta}{1+c} = \frac{mg\sin\theta}{c^{-1}+1}. \tag{12.26}$$

Note that the force of static friction plays a dual role. It decreases the center-of-mass speed and acceleration of the rolling object by reducing the magnitude of the vector sum of the forces exerted on it: $\sum F_x = F^G_{\mathrm{Eo}x} - F^s_{\mathrm{ro}}$. It also causes the torque that gives the object a rotational acceleration. In the absence of static friction, there would be no torque and objects would never roll—they would only slide.

Wheels certainly make it easier to push objects, but as Eq. 12.26 shows, they do not eliminate friction. In fact, the force of static friction to which wheels are subject slows them down, and as we have seen, the force of static friction can be *greater* than the force of kinetic friction. The next examples help clarify this paradox.



12.10 (*a*) A cylindrical shell and a solid cylinder, made of the same material and having the same inertia m and radius R, roll down a ramp. Is the force of static friction exerted on the shell greater than, smaller than, or equal to that on the solid cylinder? Explain. (*b*) For $\mu_s = 1$, what is the maximum angle of incline θ on which these objects can roll without slipping?

EXAMPLE 12.7 Bicycle friction

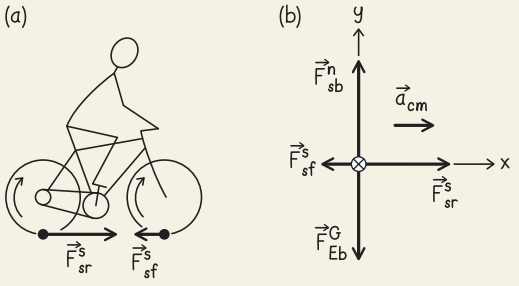
As you accelerate from rest on a bicycle, how does the magnitude of the force of friction exerted by the road surface on the rear wheel compare with the magnitude of the force of friction exerted by the road surface on the front wheel? Ignore air resistance, assume both wheels have the same inertia m_w and the same radius R, and let m_{comb} be the combined inertia of you and your bicycle (including both wheels).

❶ **GETTING STARTED** I begin by making a sketch of my bicycle (**Figure 12.40*a***). As I push on the pedals, the chain causes a torque on the rear wheel that makes the wheel rotate so that it exerts a rearward force on the road. Consequently the road surface exerts on the rear wheel a contact force in the opposite direction, and the tangential component of this force—the force of static friction—is what makes the bicycle go forward. This force, which I label $\vec{F}^s_{sr}$, is one of the forces I need for my comparison. (Note the subscripts here: r stands for "rear," not "road"; this is the frictional force exerted by the (road) surface on the rear wheel.)

The front wheel, because it is not connected to the chain, is merely pushed along with the rest of the bicycle. To make this wheel rotate in the right direction, the road surface must exert on it a rearward-pointing force of static friction $\vec{F}^s_{sf}$ (subscript f for "front" here). The torque caused by this frictional force is what makes the front wheel rotate, and this is the second force in my comparison.

Armed with this information, I make a free-body diagram for the bicycle (Figure 12.40*b*), showing the two frictional forces, the upward component of the contact force $\vec{F}^n_{sb}$ exerted by the road surface, and the downward force of gravity $\vec{F}^G_{Eb}$. Because the bicycle accelerates forward, I draw the vector arrow for $\vec{F}^s_{sr}$ longer than the vector arrow for $\vec{F}^s_{sf}$. Finally, I show the direction in which the bicycle's center of mass accelerates and choose a set of axes, letting the x axis point in the direction of motion. I also choose clockwise as the direction of increasing ϑ because that is the direction in which the wheels rotate.

Figure 12.40



❷ **DEVISE PLAN** The equation of motion in the form given by Eq. 8.45, $\vec{a}_{cm} = \Sigma\vec{F}_{ext}/m$, relates the vector sum of the forces shown in my free-body diagram to the acceleration $\vec{a}_{cm}$ of the bicycle's center of mass. Thus this equation gives me an expression that contains the two forces I'm after and seems a good place to begin. This expression contains the translational acceleration, however, whereas the variables in this problem are rotational, so I will probably need Eq. 12.23 relating α_ϑ and a_{cmx}. Focusing on the front wheel first, I can use the rotational equation of motion, Eq. 12.5, $\tau_\vartheta = I\alpha_\vartheta$, to relate the torque caused by $\vec{F}^s_{sf}$ to α_ϑ for this wheel and then use Table 11.3 to express everything in terms of the given variables m_w and R. Once I have my expression for the magnitude F^s_{sf}, I can work with Eq. 8.45 to obtain an expression for F^s_{sr} in terms of F^s_{sf} and make my comparison.

❸ **EXECUTE PLAN** The vector sum of the forces in the x direction is, from Eq. 8.45,

$$\Sigma F_x = F^s_{sr} - F^s_{sf} = m_{comb}\, a_{cmx}, \tag{1}$$

so I have an expression that contains both of the forces I'm after.

According to Eq. 12.23, the rotational acceleration of the rear wheel is $\alpha_\vartheta = a_{cmx}/R$, and the rotational acceleration of the front wheel must be the same as that of the rear wheel. To give the front wheel rotational acceleration α_ϑ, the force $\vec{F}^s_{sf}$ must cause a torque $\tau_\vartheta = I\alpha_\vartheta$ on the wheel, where I is the wheel's rotational inertia. If I treat the wheel as a thin hoop of inertia m_w, I have, from Table 11.3, $I = m_wR^2$, which means the torque on the front wheel is $\tau_\vartheta = m_wR^2\alpha_\vartheta$. The lever arm distance for this torque is equal to the radius R of the wheel, and so for this wheel $\tau_\vartheta = F^s_{sf}R = m_wR^2\alpha_\vartheta$. I know that for both wheels $\alpha_\vartheta = a_{cmx}/R$, however, so I can write

$$F^s_{sf}R = (m_wR^2)\left(\frac{a_{cmx}}{R}\right).$$

Dividing both sides by R yields one of the two forces I must compare:

$$F^s_{sf} = m_w a_{cmx}. \tag{2}$$

This is the magnitude of the force of static friction required to make the front wheel roll when the bicycle has acceleration a_{cmx}. The smaller the inertia of the wheel, the smaller this force magnitude.

Turning to the other force I need for my comparison—the force exerted on the rear wheel—I begin by eliminating a_{cmx} from Eq. 1 using Eq. 2 to get

$$F^s_{sr} - F^s_{sf} = \frac{m_{comb}}{m_w}F^s_{sf}$$

or $$F^s_{sr} = F^s_{sf}\left(1 + \frac{m_{comb}}{m_w}\right) = F^s_{sf}\left(\frac{m_{comb} + m_w}{m_w}\right).$$

The magnitude of the force of static friction exerted on the rear wheel is greater than that exerted on the front wheel by the factor $(m_{comb} + m_w)/m_w$. ✔

❹ **EVALUATE RESULT** Because $m_{comb} \gg m_w$, the magnitude of the force of static friction exerted on the front wheel is much smaller than that on the rear wheel, as I expect.

QUANTITATIVE TOOLS

12.11 (*a*) In Example 12.7, does the force of static friction exerted on the rear wheel cause a torque on the wheel? If so, in which direction? If not, why not? (*b*) In what direction is the sum of the torques on the rear wheel? What does this tell you about the relative magnitudes of the individual torques on the wheel?

12.7 Torque and energy

Torques cause objects to accelerate rotationally and thus cause a change in their rotational kinetic energy. To calculate this energy change, consider the object shown in **Figure 12.41**. A force $\vec{F}$ is exerted at point P on the object, a distance r from the axis of rotation. Let this force be such that the magnitude of its component perpendicular to r remains constant as the object rotates. This component thus causes a constant torque $\tau_\vartheta = +rF_\perp$ about the axis. Equation 12.10 relates the sum of the torques to the rotational acceleration. By applying the chain rule (Appendix B), we can write this equation in the form

$$\Sigma\tau_{\text{ext}\,\vartheta} = I\alpha_\vartheta = I\frac{d\omega_\vartheta}{dt} = I\frac{d\omega_\vartheta}{d\vartheta}\frac{d\vartheta}{dt} = I\frac{d\omega_\vartheta}{d\vartheta}\omega_\vartheta, \tag{12.27}$$

so

$$\left(\Sigma\tau_{\text{ext}\,\vartheta}\right)d\vartheta = I\omega_\vartheta\, d\omega_\vartheta. \tag{12.28}$$

Integrating the left side of this equation yields

$$\int_{\vartheta_\text{i}}^{\vartheta_\text{f}}\left(\Sigma\tau_{\text{ext}\,\vartheta}\right)d\vartheta = \left(\Sigma\tau_{\text{ext}\,\vartheta}\right)\int_{\vartheta_\text{i}}^{\vartheta_\text{f}} d\vartheta = \left(\Sigma\tau_{\text{ext}\,\vartheta}\right)\Delta\vartheta. \tag{12.29}$$

Integrating the right side after pulling the (constant) rotational inertia I out of the integral gives

$$\int_{\omega_{\vartheta,\text{i}}}^{\omega_{\vartheta,\text{f}}} I\omega_\vartheta\, d\omega_\vartheta = I\int_{\omega_{\vartheta,\text{i}}}^{\omega_{\vartheta,\text{f}}}\omega_\vartheta\, d\omega_\vartheta = I\left[\tfrac{1}{2}\omega_\vartheta^2\right]_{\omega_{\vartheta,\text{i}}}^{\omega_{\vartheta,\text{f}}}$$

$$= \tfrac{1}{2}I\omega_{\vartheta,\text{f}}^2 - \tfrac{1}{2}I\omega_{\vartheta,\text{i}}^2 = K_{\text{rot,f}} - K_{\text{rot,i}}. \tag{12.30}$$

Equations 12.28–12.30 thus show that the change in the rotational kinetic energy, $\Delta K_\text{rot} \equiv K_{\text{rot,f}} - K_{\text{rot,i}}$, is

$$\Delta K_\text{rot} = \left(\Sigma\tau_{\text{ext}\,\vartheta}\right)\Delta\vartheta \quad \text{(constant torques, rigid object)}. \tag{12.31}$$

This equation states that the change in an object's rotational kinetic energy is equal to the product of the sum of the torques on the object and the object's rotational displacement. Note the analogy to Eq. 9.14, which relates the change in an object's center-of-mass kinetic energy to the product of the vector sum of the forces exerted on it and the object's displacement:

$$\Delta K_\text{cm} = \left(\Sigma F_\text{ext}\right)\Delta x_\text{cm} \quad \text{(constant forces, one dimension)}. \tag{12.32}$$

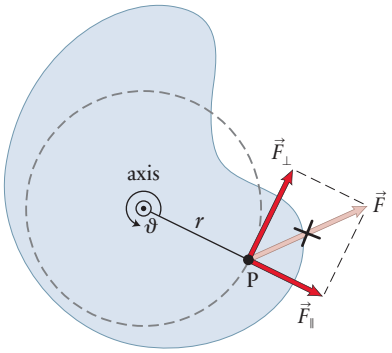
The kinetic energy of an object or system that is in both translational and rotational motion is equal to the sum of its center-of-mass and rotational kinetic energies:

$$K = K_\text{cm} + K_\text{rot} = \tfrac{1}{2}mv_\text{cm}^2 + \tfrac{1}{2}I\omega^2, \tag{12.33}$$

and the change in this kinetic energy is given by the sum of the changes in Eqs. 12.31 and 12.32:

$$\Delta K = \Delta K_\text{cm} + \Delta K_\text{rot}. \tag{12.34}$$

Figure 12.41 A rigid object subject to a constant torque caused by a force $\vec{F}$ exerted on it undergoes a rotational displacement $\Delta\vartheta$.



EXAMPLE 12.8 Rolling down a ramp

A solid cylindrical object of inertia m, rotational inertia I, and radius R rolls down a ramp that makes an angle θ with the horizontal. By how much does the cylinder's energy increase if it is released from rest and its center of mass drops a vertical distance h?

❶ **GETTING STARTED** I am given information about an object in the shape of a solid cylinder—inertia, rotational inertia, radius, and initial speed—and my task is to find out how much the object's energy has increased once it has rolled down a ramp such that its center of mass has traveled a vertical distance h. The object accelerates down the incline under the influence of the force of gravity. I therefore begin by making a sketch of the situation and drawing both free-body and extended free-body diagrams (**Figure 12.42**). The object is subject to a gravitational force exerted by Earth and a contact force exerted by the ramp. If I choose my axes as shown in my sketch, the contact force exerted by the ramp has a normal component $\vec{F}^n_{ro}$ in the y direction and a tangential component $\vec{F}^s_{ro}$ in the negative x direction due to static friction.

❷ **DEVISE PLAN** As the object rolls, both its translational and rotational kinetic energies increase. Because the shape of the object does not change and because static friction is nondissipative, the object's internal energy does not change. I can use Eq. 12.32 for the change in translational kinetic energy and Eq. 12.31 for the change in rotational kinetic energy. To express the two factors on the right in Eq. 12.32 in terms of my given variables, I use the geometry of the situation to express both factors in terms of $\sin\theta$.

❸ **EXECUTE PLAN** The change in translational kinetic energy is given by Eq. 12.32, $\Delta K_{cm} = \left(\sum F_{ext\,x}\right)\Delta x_{cm}$, and the vector sum of the forces exerted on the object in the x direction is

$$\sum F_{ext\,x} = mg\sin\theta - F^s_{ro}$$

or, using Eq. 12.26,

$$\sum F_{ext\,x} = +mg\sin\theta\left(1 - \frac{1}{1+c^{-1}}\right),$$

where c is the object's shape factor. Because the displacement of the object's center of mass along the plane is $\Delta x_{cm} = h/\sin\theta$, the change in its translational kinetic energy is

$$\Delta K_{cm} = \left(\sum F_{ext\,x}\right)\Delta x_{cm} = mgh\left(1 - \frac{1}{1+c^{-1}}\right).$$

Next, I use Eq. 12.31 to calculate the change in the object's rotational kinetic energy. From my extended free-body diagram I see that only the force of static friction causes a (positive) torque, so $\sum\tau_{ext\,\vartheta} = +F^s_{ro}R$. I can find the object's rotational displacement $\Delta\vartheta$ from Eq. 12.18:

$$\Delta\vartheta = \frac{\Delta x_{cm}}{R} = \left(\frac{h}{\sin\theta}\right)\left(\frac{1}{R}\right) = +\frac{h}{R\sin\theta},$$

so, from Eq. 12.31,

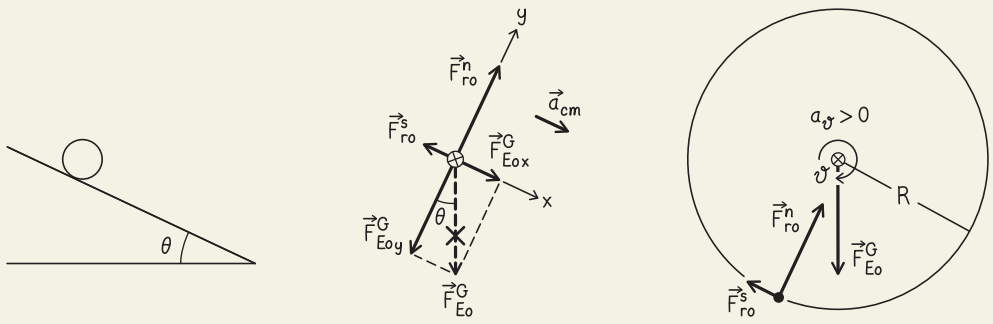
$$\Delta K_{rot} = (F^s_{ro}R)\left(\frac{h}{R\sin\theta}\right) = \left(\frac{mg\sin\theta}{1+c^{-1}}\right)\left(\frac{h}{\sin\theta}\right) = \frac{mgh}{1+c^{-1}},$$

where I have again used Eq. 12.26 to substitute for F^s_{ro}. Adding the two changes in kinetic energy, I obtain

$$\Delta E = \Delta K_{cm} + \Delta K_{rot}$$
$$= mgh\left(1 - \frac{1}{1+c^{-1}}\right) + mgh\left(\frac{1}{1+c^{-1}}\right) = mgh. \checkmark$$

❹ **EVALUATE RESULT** My result indicates that the object's energy changes by the same amount it would change if it were simply in free fall! In other words, the only work done on the object is the work done by the gravitational force: $\vec{F}^G_{Eo}\cdot\Delta\vec{r} = mgh$ (see Section 10.9). This implies that the work done by all other forces on the object is zero. The normal force does no work on the object because it is perpendicular to the displacement of the object, but what is the work done by the force of static friction on the object? The object's displacement, $h/\sin\theta$, lies along the line of action of the force of static friction, and so it is tempting to write $-F^s_{ro}(h/\sin\theta)$ for the work done by the force of static friction on the object. However, the point of application for $\vec{F}^s_{ro}$ has zero velocity. At each instant, a different point on the object's surface touches the ramp, but the instantaneous velocity of that point is zero. The force displacement for $\vec{F}^s_{ro}$ is thus zero, so the work done by this force on the object is zero as well.

Figure 12.42



QUANTITATIVE TOOLS

Example 12.8 shows that *the force of static friction does no work on a rolling object*! The only thing it does is take away some of the translational kinetic energy and generate an equal amount of rotational kinetic energy. It takes away translational kinetic energy by reducing the magnitude of the vector sum of the forces and converts it to rotational kinetic energy by causing a torque.

12.12 In Figure 12.41, does the force exerted on the rigid object do work on the object?

12.8 The vector product

To complete our treatment of rotation, we must examine the vectorial nature of the rotational components. As we saw in the first part of this chapter, an object's rotational velocity can be described by a vector $\vec{\omega}$. Angular momentum, which is the product of a scalar and $\vec{\omega}$, $\vec{L} = I\vec{\omega}$, is therefore also a vector and points in the same direction as $\vec{\omega}$. Because the rotational acceleration $\vec{\alpha}$ is the time derivative of $\vec{\omega}$, it must be a vector, too. Torque, which is the product of a scalar and $\vec{\alpha}$, $\vec{\tau} = I\vec{\alpha}$, therefore points in the same direction as $\vec{\alpha}$.

Consider the object in **Figure 12.43**. A force $\vec{F}$ exerted at point P increases the magnitude of the object's rotational velocity about the rotation axis at O. As the magnitude of the rotational velocity increases, the length of the vector representing $\vec{\omega}$ increases. Because the rotation takes place in a plane, the direction of $\vec{\omega}$ remains along the axis of rotation. The direction of $\vec{\alpha}$ is related to the direction of the vector representing the change in rotational velocity $\Delta\vec{\omega}$. Because the direction of $\vec{\omega}$ remains along the axis of rotation, $\Delta\vec{\omega}$, $\vec{\alpha}$, and $\vec{\tau}$ also lie along the rotation axis.

The magnitude of the torque is $rF \sin\theta$, where θ is the angle between $\vec{r}$ and $\vec{F}$. In Section 10.9, I showed that work on a system can be written as a scalar product of two vectors. To account for the vectorial nature of torque, we must introduce a new kind of product of two vectors: one that generates a new vector. Specifically, this product of the vectors $\vec{r}$ and $\vec{F}$ must generate a vector whose magnitude is $rF \sin\theta$ and whose direction is as indicated in **Figure 12.44**. Such a product is called the **vector product.** The vector product of two vectors $\vec{A}$ and $\vec{B}$, written $\vec{A} \times \vec{B}$, is a vector whose magnitude is

$$|\vec{A} \times \vec{B}| \equiv AB \sin\theta, \tag{12.35}$$

where θ is the angle between $\vec{A}$ and $\vec{B}$ when they are placed tail to tail (and $\theta \leq 180°$). The direction of the vector $\vec{A} \times \vec{B}$ is determined as illustrated in Figure 12.44*a*: Place vectors $\vec{A}$ and $\vec{B}$ with their tails together and apply the right-hand rule to the direction of rotation obtained by rotating $\vec{A}$ into $\vec{B}$

Figure 12.43 For a rotation about a fixed axis with increasing rotational speed, the rotational velocity, rotational acceleration, and torque vectors all point along the rotation axis.

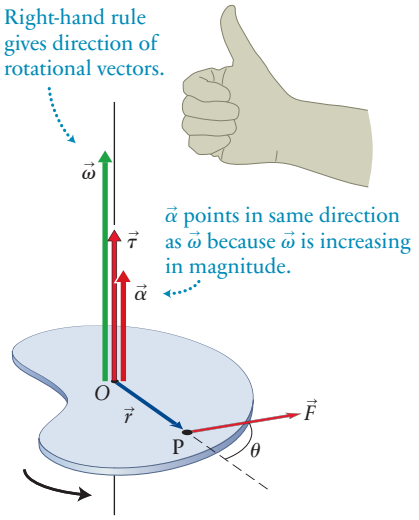
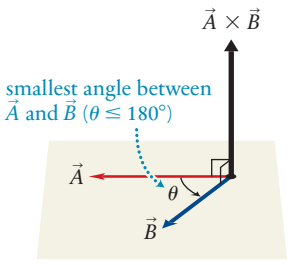
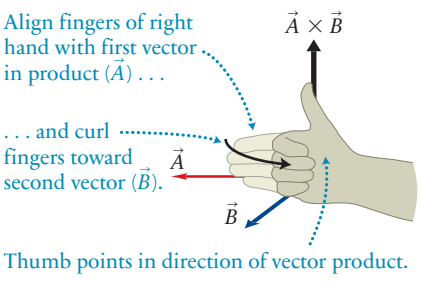


Figure 12.44 The vector product of two vectors.

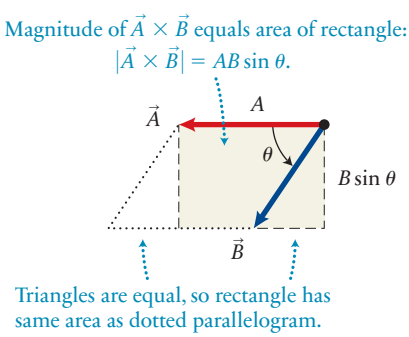
(*a*) Vector product $\vec{A} \times \vec{B}$



(*b*) Finding the direction of a vector product



(*c*) Magnitude of a vector product



through the smallest angle. When you point the fingers of your right hand along the direction of the first vector in the product and curl them toward the direction of the second vector, your thumb points in the direction of the vector product (Figure 12.44*b*).

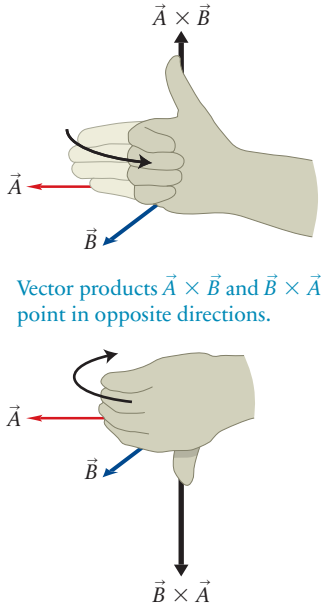
Figure 12.44*c* shows a geometric interpretation of the magnitude of the vector product. The area of the shaded rectangle is equal to $AB \sin\theta$, and because the two triangles are identical, the area of the parallelogram defined by the vectors $\vec{A}$ and $\vec{B}$ is equal to the area of the shaded rectangle. Therefore:

The magnitude of the vector product of two vectors is equal to the area of the parallelogram defined by them.

(The vector product need not have units of area, however.) The magnitude of the vector product is greatest when the two vectors are perpendicular to each other and $\sin\theta = \sin 90° = 1$. When the two vectors are parallel or antiparallel, the parallelogram is reduced to a straight line and the vector product is zero. In particular, the vector product of a vector with itself is zero:

$$\vec{A} \times \vec{A} = \vec{0}. \tag{12.36}$$

Figure 12.45



Vector products $\vec{A} \times \vec{B}$ and $\vec{B} \times \vec{A}$ point in opposite directions.

Another interesting property of the vector product is that the multiplication sequence is important. The vector product $\vec{B} \times \vec{A}$ is equal in magnitude to $\vec{A} \times \vec{B}$ but points in the opposite direction, as shown in **Figure 12.45**:

$$\vec{B} \times \vec{A} = -(\vec{A} \times \vec{B}). \tag{12.37}$$

Using the definition of vector product, we can write the torque about the origin caused by a force $\vec{F}$ as

$$\vec{\tau} = \vec{r} \times \vec{F}. \tag{12.38}$$

You can easily verify that this product has the correct magnitude and direction.

Because torque is the time derivative of angular momentum (Eq. 12.12) and force is the time derivative of momentum (Eq. 8.4), we can write Eq. 12.38 as

$$\frac{d\vec{L}}{dt} = \vec{r} \times \frac{d\vec{p}}{dt}. \tag{12.39}$$

To see what this expression means, we use the product rule (Appendix B) to find the time derivative of the vector product $\vec{r} \times \vec{p}$:

$$\frac{d}{dt}(\vec{r} \times \vec{p}) = \frac{d\vec{r}}{dt} \times \vec{p} + \vec{r} \times \frac{d\vec{p}}{dt}. \tag{12.40}$$

The first term on the right side vanishes because the vector product of a vector and itself is zero:

$$\frac{d\vec{r}}{dt} \times \vec{p} = \vec{v} \times (m\vec{v}) = m(\vec{v} \times \vec{v}) = 0. \tag{12.41}$$

Therefore the right side of Eq. 12.39 is equal to the time derivative of the vector product $\vec{r} \times \vec{p}$:

$$\frac{d\vec{L}}{dt} = \frac{d}{dt}(\vec{r} \times \vec{p}). \tag{12.42}$$

The angular momentum vector can therefore be written as

$$\vec{L} = \vec{r} \times \vec{p}. \tag{12.43}$$

Like the rotational velocity vector, $\vec{L}$ is perpendicular to the plane defined by $\vec{r}$ and $\vec{p}$, a direction in which nothing moves—there is certainly no momentum in this direction.

Because it applies to any type of motion and not just rotational motion, the expression $\vec{L} = \vec{r} \times \vec{p}$ is much more general than $\vec{L} = I\vec{\omega}$. Consider, for example, the motion of the isolated particle in **Figure 12.46**. Because the particle is isolated, its trajectory is a straight line and its momentum is constant. From Eq. 12.35 we see that the magnitude of the particle's angular momentum with respect to the origin O is

$$L = r\,mv \sin\theta = r_\perp mv, \tag{12.44}$$

where θ is the angle between $\vec{r}$ and $\vec{p}$ and $r_\perp$ is the lever arm distance of $\vec{p}$ relative to O. This is the same result we obtained in Chapter 11 (Eq. 11.36). The right side of Eq. 12.44 is indeed constant. Even though $\vec{r}$ changes in direction and magnitude, the lever arm distance $r_\perp$ remains constant. Consequently the product $\vec{r} \times \vec{p}$, perpendicular to the plane defined by O and the particle's trajectory, is unchanged in direction and magnitude.

Figure 12.46 The angular momentum of an isolated particle about an arbitrary point (called origin O here) is constant because $r \sin\theta = r_\perp$ is the same regardless of the position of the particle.

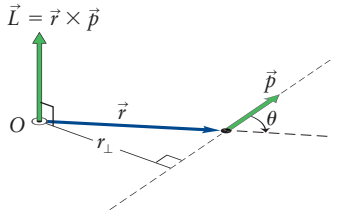


Table 12.1 Translational and rotational dynamics

	Translation	Rotation	Relationship
Momentum	$\vec{p} \equiv m\vec{v}$	$\vec{L} \equiv I\vec{\omega}$	$\vec{L} = \vec{r} \times \vec{p}$
Momentum change	$\Delta\vec{p} = \Sigma\vec{F}_{\text{ext}}\Delta t$	$\Delta\vec{L} = \Sigma\vec{\tau}_{\text{ext}}\Delta t$	
Force	$\Sigma\vec{F}_{\text{ext}} \equiv \dfrac{d\vec{p}}{dt} = m\vec{a}_{\text{cm}}$	$\Sigma\vec{\tau}_{\text{ext}} = \dfrac{d\vec{L}}{dt} = I\vec{\alpha}$	$\vec{\tau} = \vec{r} \times \vec{F}$
Kinetic energy	$K_{\text{cm}} = \frac{1}{2}mv^2$	$K_{\text{rot}} = \frac{1}{2}I\omega^2$	
Kinetic energy change	$\Delta K_{\text{cm}} = \left(\Sigma\vec{F}_{\text{ext}}\right)\Delta\vec{r}_{\text{cm}}$	$\Delta K_{\text{rot}} = \left(\Sigma\tau_{\text{ext}\,\vartheta}\right)\Delta\vartheta$	

Table 12.1 summarizes the results obtained in this chapter and the remarkable parallels between the equations that describe the dynamics of translational and rotational motion. The two sets of equations are mathematically identical given the correspondences in **Table 12.2**. The only new mathematical tool we have introduced in this chapter is the vector product. Apart from this product, which is required to account for the vectorial nature of rotation, the mathematical framework for the two types of motion is identical. Finally, note that—unlike their translational counterparts—rotational inertia, angular momentum, and torque depend on the choice of origin and axis. You should therefore always specify your choice of axis when dealing with these quantities and—most important—consistently use the same axis for the different quantities throughout your calculations.

Table 12.2 Corresponding translational and rotational quantities

position	$\vec{r}$	$\leftrightarrow$	ϑ
velocity	$\vec{v}$	$\leftrightarrow$	$\vec{\omega}$
acceleration	$\vec{a}$	$\leftrightarrow$	$\vec{\alpha}$
inertia	m	$\leftrightarrow$	I

12.13 Consider a particle moving at constant speed along a circular trajectory centered on the origin. Write the relationship of the particle's velocity $\vec{v}$, rotational velocity $\vec{\omega}$, and position vector $\vec{r}$ in the form of a vector product.

Chapter Summary

Torque (Sections 12.1, 12.5)

Concepts The **torque** due to a force exerted on an object is the tendency of the force to give rotational acceleration to the object. The torque about some pivot is the product of the magnitude of the force and the lever arm distance. The SI units for torque are $\mathrm{N \cdot m} = \mathrm{kg \cdot m^2/s^2}$.

An object is in *translational equilibrium* if the vector sum of all the external forces exerted on it is zero. An object is in **rotational equilibrium** if the vector sum of the external torques is zero. If the vector sums of both the forces and the torques are zero, the object is in **mechanical equilibrium.**

Quantitative Tools If $\vec{r}$ is the position vector from a pivot to the location at which a force $\vec{F}$ is exerted on an object and θ is the angle between $\vec{r}$ and $\vec{F}$, the torque τ produced by the force about the pivot is

$$\tau \equiv rF \sin\theta = r_\perp F = rF_\perp, \tag{12.1}$$

where $r_\perp$ is the component of $\vec{r}$ perpendicular to $\vec{F}$ and $F_\perp$ is the component of $\vec{F}$ perpendicular to $\vec{r}$.

Translational equilibrium: $\Sigma\vec{F}_{\mathrm{ext}} = \vec{0}$.

Rotational equilibrium: $\Sigma\tau_{\mathrm{ext}\,\vartheta} = 0$.

Mechanical equilibrium:

$$\Sigma\tau_{\mathrm{ext}\,\vartheta} = 0 \quad \text{and} \quad \Sigma\vec{F}_{\mathrm{ext}} = \vec{0}. \tag{12.14}$$

Rotation of a rigid object (Sections 12.1, 12.2, 12.5–12.7)

Concepts In **free rotation,** an object rotates about its center of mass. The translational and rotational motions are analyzed separately, and the torques and rotational inertia are computed about the center of mass. In *rotation about a fixed axis*, an object is constrained to rotate about a physical axis. Torques and rotational inertia are computed about that axis.

When an object is **rolling without slipping** across a surface, there is no relative motion at the location where the object touches the surface. The object rotates about its geometric center as it translates across the surface. Torques and rotational inertia are computed about the object's center of mass, which for a symmetrical object is the geometric center.

The external force exerted on a rolling object changes the object's center-of-mass kinetic energy. The external torque on a rolling object changes the object's rotational kinetic energy.

Quantitative Tools For both particles and extended bodies, the vector sum of the torques when rotation is about a fixed axis is

$$\Sigma\tau_{\mathrm{ext}\,\vartheta} = I\alpha_\vartheta. \tag{12.10}$$

For an object of radius R that is rolling without slipping, the motion of the center of mass is described by

$$v_{\mathrm{cm}x} = R\omega_\vartheta \tag{12.19}$$

$$a_{\mathrm{cm}x} = R\alpha_\vartheta \tag{12.23}$$

$$\Sigma F_{\mathrm{ext}x} = ma_{\mathrm{cm}x}$$

$$\Sigma\tau_{\mathrm{ext}\,\vartheta} = I\alpha_\vartheta. \tag{12.10}$$

The change in an object's rotational kinetic energy resulting from torques is

$$\Delta K_{\mathrm{rot}} = \left(\Sigma\tau_{\mathrm{ext}\,\vartheta}\right)\Delta\vartheta \quad \text{(constant torques, rigid object)}. \tag{12.31}$$

The kinetic energy of a rolling object is

$$K = K_{\mathrm{cm}} + K_{\mathrm{rot}} = \tfrac{1}{2}mv_{\mathrm{cm}}^2 + \tfrac{1}{2}I\omega^2 \tag{12.33}$$

and the change in the kinetic energy of a rolling object is

$$\Delta K = \Delta K_{\mathrm{cm}} + \Delta K_{\mathrm{rot}}. \tag{12.34}$$

Angular momentum (Sections 12.1, 12.5)

Concepts A **rotational impulse** J_ϑ is the amount of angular momentum transferred to a system from the environment by external torques.

If the sum of the external torques due to forces exerted on a system is zero, the angular momentum of the system remains constant.

Quantitative Tools External torque caused by forces exerted on an object causes the object's angular momentum L_ϑ to change:

$$\Sigma\tau_{\mathrm{ext}\,\vartheta} = \frac{dL_\vartheta}{dt}. \tag{12.12}$$

The **angular momentum law** says that the change in the angular momentum of an object is equal to the **rotational impulse** given to the system:

$$\Delta L_\vartheta = J_\vartheta. \tag{12.15}$$

If the constant external torques on a system last for a time interval Δt, the **rotational impulse equation** says that the rotational impulse is

$$J_\vartheta = \left(\Sigma \tau_{\text{ext}\,\vartheta}\right)\Delta t. \tag{12.17}$$

The law of conservation of angular momentum states that if $\Sigma \tau_{\text{ext}\,\vartheta} = 0$, then $dL_\vartheta/dt = 0$. This means that

$$\Sigma \tau_{\text{ext}\,\vartheta} = \frac{dL_\vartheta}{dt} = 0 \Rightarrow \Delta L_\vartheta = 0. \tag{12.13}$$

Rotational quantities as vectors (Sections 12.4, 12.8)

Concepts A **polar vector** is a vector associated with a displacement. An **axial vector** is a vector associated with a rotation direction. This vector points along the rotation axis, and its direction is given by the **right-hand rule** for axial vectors: When you curl the fingers of your right hand along the direction of rotation, your outstretched thumb points in the direction of the vector that specifies that rotation.

Right-hand rule for vector products: When you align the fingers of your right hand along the first vector in a vector product and curl them from that vector to the second vector in the product through the smaller angle between the vectors, your outstretched thumb points in the direction of the vector product.

The magnitude of the vector product of two vectors is equal to the area of the parallelogram defined by them.

Quantitative Tools The magnitude of the **vector product** of vectors $\vec{A}$ and $\vec{B}$ that make an angle $\theta \leq 180°$ between them when they are tail to tail is

$$|\vec{A} \times \vec{B}| = AB \sin\theta. \tag{12.35}$$

If $\vec{r}$ is the vector from the origin of a coordinate system to the location where a force $\vec{F}$ is exerted, the torque about the origin due to $\vec{F}$ is

$$\vec{\tau} = \vec{r} \times \vec{F}. \tag{12.38}$$

If $\vec{r}$ is the vector from the origin of a coordinate system to a particle that has momentum $\vec{p}$, the angular momentum of the particle about the origin is

$$\vec{L} = \vec{r} \times \vec{p}. \tag{12.43}$$

Questions and Problems

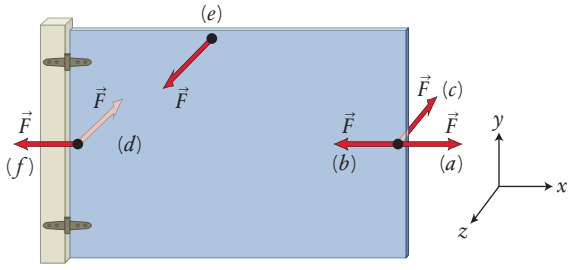
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

12.1 Torque and angular momentum

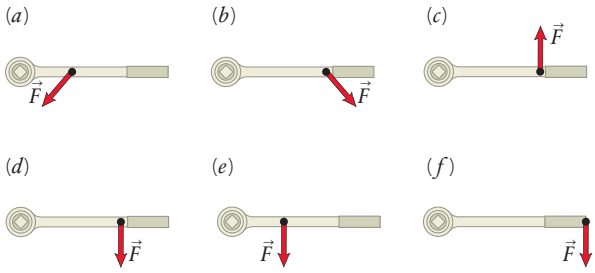
1. To open a stuck jar lid, wearing rubber gloves sometimes works. Why does that help? •
2. You are using a steel crowbar to lift a big rock, with a smaller stone as a fulcrum at the center of the crowbar, but you are not quite able to lift the rock. What can you change about this setup to lift the rock? •
3. The specifications for tightening the bolts on a car engine's valve cover are given in terms of torque rather than in terms of how much force should be exerted on the tightening wrench. Why is torque a better specification here? •
4. When the wrench you are working with does not loosen a nut, you can sometimes succeed by slipping a length of pipe over the end of the wrench and pushing at the end of the pipe. Why does this work? ••
5. In which cases illustrated in Figure P12.5 does pushing in the way and at the location shown cause a torque about a vertical axis passing through the hinges? The dot shows the point of application of the force in each case, and the dashed lines help indicate the line of action of each force. Forces (*a*), (*b*), and (*f*) are exerted parallel to the gate; forces (*c*), (*d*), and (*e*) are exerted perpendicular to the gate. ••

Figure P12.5



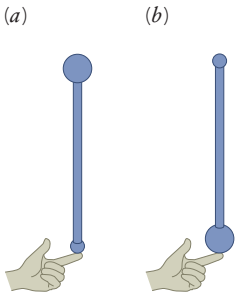
6. Rank, from greatest to smallest, the magnitudes of the torques about the wrench head in Figure P12.6 caused by exerting the same force at the different positions shown. ••

Figure P12.6



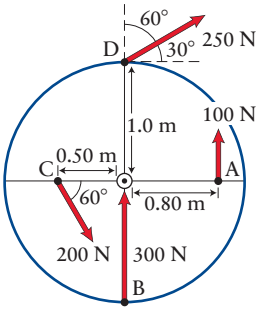
7. Which way is easier to balance the asymmetrical baton of Figure P12.7 on your finger? Explain your choice. ••

Figure P12.7



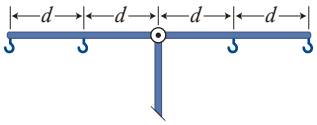
8. Explain why, when a car makes a sharp turn on an unbanked road, the wheels on the inside of the turn tend to come off the ground. ••
9. Figure P12.9 is a top view of a hand-driven carousel in a park. Four parents of the children riding on the device each push on the carousel in different ways. The forces exerted by each of the four parents and the positions at which the forces are exerted are labeled A–D. Rank the parents' pushes in order of increasing magnitude of the torque they cause on the carousel. ••

Figure P12.9



10. The uniform rod shown in Figure P12.10 is fixed to an axis going through its center, so that it is free to rotate clockwise or counterclockwise. It has four hooks: one at each end of the rod and one halfway between the center (axis) and each end. You need to hang four objects from this rod, with one and only one object on each hook, after which the rod must be balanced. The objects have masses M, $3M$, $5M$, and $9M$, respectively. How can this be accomplished? •••

Figure P12.10



12.2 Free rotation

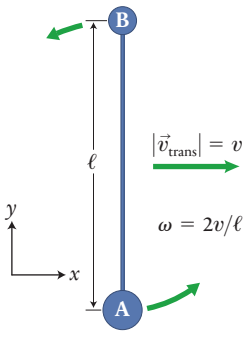
11. A very thin bar of length ℓ and negligible inertia connects two blocks A and B that have inertias $4m$ and m, respectively (Figure P12.11). When this rod-block system is made to spin with no translational motion and no mechanical pivot, each block moves through a circular path in space, with circumferences C_A and C_B. Determine the ratio C_A/C_B. ••

Figure P12.11



12. The system shown in Figure P12.12 consists of two balls A and B connected by a thin rod of negligible mass. Ball A has three times the inertia of ball B and the distance between the two balls is ℓ. If the system has a translational velocity of v in the x direction and is spinning counterclockwise at an angular speed of $\omega = 2v/\ell$, determine the ratio of the instantaneous speeds of the two balls v_A/v_B at the moment shown. ••

Figure P12.12



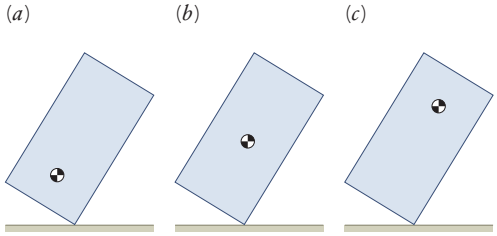
13. Put your index fingers about 1 m apart and rest a meterstick or similar object on top of them. Then, while holding the meterstick parallel to the floor, slowly draw your fingers toward each other. You will find that first one finger slips easily under the stick, then the other finger, then the first one again, then the other one again. No matter how much alternating takes place, however, your fingers *always* meet at the center of the stick. Explain what is happening. •••

12.3 Extended free-body diagrams

14. To change a light bulb, you climb halfway up a stepladder. Draw an extended free-body diagram for the ladder. •

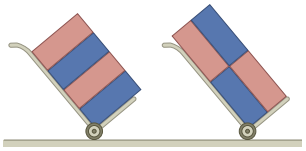
15. Which refrigerator in Figure P12.15 is in the greatest danger of tipping over? Note positions of center of mass. •

Figure P12.15



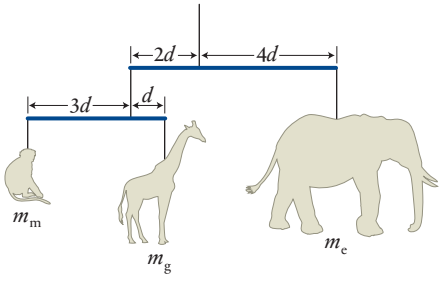
16. Which configuration of the four uniform boxes in Figure P12.16 makes the hand-truck easier to handle? Explain your answer. ••

Figure P12.16



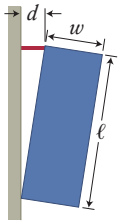
17. A father and son are carrying a sofa down the stairs, with the son leading. Consider the sofa's center of mass to be at the geometric center. (*a*) Draw an extended free-body diagram for the sofa. (*b*) Who carries the greater share of the load? (*c*) Would your answer to part *b* be the same if the men were carrying instead a sheet of plywood, with its flat sides parallel to the stairs? ••
18. In the mobile in Figure P12.18, what are the inertias of the giraffe and of the elephant in terms of the inertia m_m of the monkey? ••

Figure P12.18



19. A painter leans a ladder at some angle against a smooth wall, with the foot of the ladder resting on wall-to-wall carpeting. Why might the ladder slip as he climbs it? ••
20. You are pushing a cart across the room, and the cart has wheels at the front and the back. Your hands are placed on top of the cart at the center (left to right) of the top edge, pushing horizontally. There is friction between the wheels and the floor. Is the normal force between the floor and the front wheels greater than, smaller than, or equal to the normal force between the floor and the rear wheels? Explain your answer using an extended free-body diagram. ••
21. The top of a ladder of inertia m rests against a smooth wall, and the foot rests on the ground. The coefficient of static friction between ground and ladder is μ_s. What is the smallest angle between the ground and the ladder such that the ladder does not slip? •••
22. A square clock of inertia m is hung on a nail driven into a wall (Figure P12.22). The length of each side of the square is ℓ, the thickness is w, and the top back edge of the clock is a distance d from the wall. Assume that the wall is smooth and that the center of mass of the clock is at the geometric center. Obtain an expression for the magnitude of the normal force exerted by the wall on the clock. •••

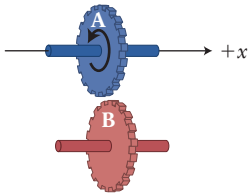
Figure P12.22



12.4 The vectorial nature of rotation

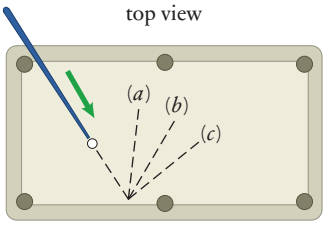
23. Beginning mechanics sometimes are taught a rule for tightening and untightening bolts: Righty-tighty, lefty-loosey. Explain why this memory device works. •
24. What is the direction of the rotational velocity vector of the jar lid when you open a jar of pickles? •
25. Gear A in Figure P12.25 is spinning such that its rotational velocity vector points in the negative x direction. It is lowered so that its edge comes in contact with the edge of gear B. What is the direction of the rotational velocity caused by the force that A exerts on B? •

Figure P12.25



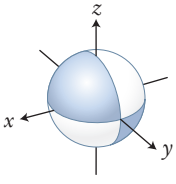
26. An experienced pool player can control the placement of a ball after it hits an edge bumper by putting spin, or "English," on the ball. This is accomplished by aiming the cue stick so that it hits the ball either left or right of the ball's center. Suppose a player gives a ball English by hitting it to the left of center (left from his perspective). Which path in Figure P12.26 does the ball take after it collides with the bumper? Account for this behavior. ••

Figure P12.26



27. The beach ball shown in Figure P12.27 undergoes a rotation of 90° around the x axis and then a rotation of 90° around the y axis. A second, identically colored ball undergoes first the rotation of 90° around the y axis and then the 90° rotation around the x axis. What additional rotations are necessary to put the second ball in exactly the same orientation as the first? ••

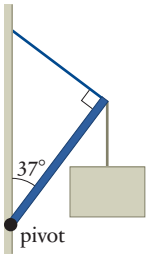
Figure P12.27



12.5 Conservation of angular momentum

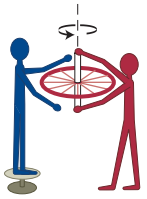
28. Why do quarterbacks throw the football with significant spin about its long axis? •
29. A 1.6-m uniform rod is being used to balance two buckets of paint, each of inertia m, one at each end of the rod. (*a*) Where is the pivot located? (*b*) If paint is removed from one bucket until its inertia is $m/3$, where must the pivot now be placed in order to keep the rod balanced? Ignore the inertia of the rod. •
30. A 0.17-kg turntable of radius 0.15 m spins about a vertical axis through its center. A constant rotational acceleration causes the turntable to accelerate from 0 to 33 revolutions per second in 9.0 s. Calculate (*a*) the rotational acceleration and (*b*) the torque required to cause this acceleration. •
31. A 44-kg child runs at 3.0 m/s tangent to a stationary 180-kg playground merry-go-round that has a radius of 1.2 m. The child then jumps on and grabs hold, causing the merry-go-round to rotate. Determine the resulting rotational velocity of the merry-go-round. ••
32. Standing on the edge of a rotating playground carousel, you take sight on the center of the carousel and walk directly toward that position. (*a*) What happens to the rotational speed of the carousel as you walk? (*b*) In the system consisting of only the carousel (a nonisolated system), any change in angular momentum is due to a torque. Because the hub of the carousel is well lubricated, the only candidate for what causes this torque is the force of friction exerted by your feet on the carousel. "But," you say, "I walked radially inward. How could I possibly cause a torque with my feet?" What's going on? ••
33. Standing with your toes touching the base of a wall, press your nose to the wall and try to stand on tiptoe. What happens and why? ••
34. Draw the angular momentum vector for a tetherball, taking as your origin the center of the pole. Is the angular momentum constant? If not, what force causes the torque that changes the angular momentum? ••
35. Imagine that an asteroid 1 km in diameter collides with Earth. Estimate the maximum fractional change in the length of the day due to this collision. ••
36. You want to hang a 10-kg sign that advertises your new business. To do this, you use a pivot to attach the base of a 5.0-kg beam to a wall (Figure P12.36). You then attach a cable to the beam and to the wall in such a way that the cable and beam are perpendicular to each other. The beam is 2.0 m long and makes an angle of 37° with the vertical. You hang the sign from the end of the beam to which the cable is attached. (*a*) What must be the minimum tensile strength of the cable (the minimum amount of tension it can sustain) if it is not to snap? (*b*) Determine the horizontal and vertical components of the force exerted by the pivot on the beam. ••

Figure P12.36



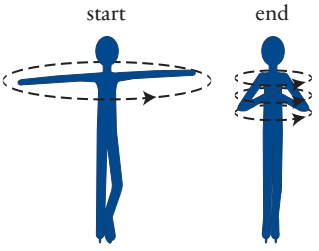
37. The top end of a two-by-four piece of lumber that is 3.0 m long is leaned at a height of 1.8 m against a smooth wall so that the bottom end makes an angle of 37° with the floor. If this board has an inertia of 4.6 kg, what are (*a*) the normal force exerted by the floor on it and (*b*) the normal force exerted by the wall on it? ••
38. Your physics instructor calls you up to the front of the class to be a demonstration assistant. She has you stand on a turntable that is free to rotate and hands you a spinning bicycle wheel, as shown in Figure P12.38. Letting go of the top of the wheel's axle, you then stop its spinning with your free hand. What happens when you stop the wheel? ••

Figure P12.38



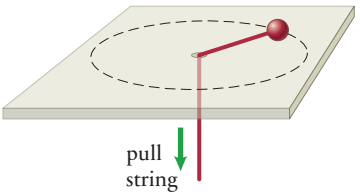
39. Helicopters have a small tail rotor as well as the large main rotor. Why is the tail rotor needed? ••
40. A 35-kg child stands on the edge of a 400-kg playground merry-go-round that is turning at the rate of 1 rev every 2.2 s. He then walks to the center of the platform. If the radius of the platform is 1.5 m, what is its rotational speed once he arrives at the center? ••
41. An ice figure skater starts out spinning at 0.85 revolution per second with her arms outstretched, as shown in Figure P12.41. She wears lightly weighted bracelets to enhance the spin-up effect when she pulls her arms in. (*a*) Calculate her final rotational speed if her rotational inertia is 3.6 kg·m^2 with her arms outstretched and 1.1 kg·m^2 with her arms pulled close to her body. (*b*) Determine the increase in her rotational kinetic energy. (*c*) Where does this added energy come from? ••

Figure P12.41



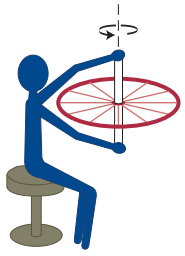
42. A 25-kg child starts at the center of a playground merry-go-round that has a radius of 2.0 m and rotational inertia of 500 kg·m^2 and walks out to the edge. If the merry-go-round has a rotational speed of 0.20 s^{-1} when she is at the center, what is its rotational speed when she gets to the edge? ••
43. In the setup in Figure P12.43, does pulling the string downward while the ball is revolving make the following properties of the ball increase, decrease, or stay the same: (*a*) angular momentum, (*b*) rotational speed, (*c*) rotational kinetic energy? ••

Figure P12.43



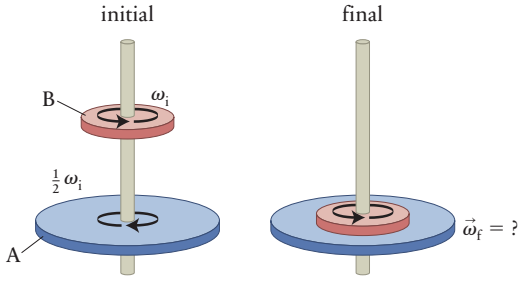
44. Your physics instructor has you sit on a chair that is free to rotate and hands you a spinning bicycle wheel (Figure P12.44). What happens when you flip the wheel over so that the end of the axle that initially pointed up now points down? ●●

Figure P12.44



45. A putty ball of inertia m moving at speed v_i slides on a slick horizontal surface along a path that is perpendicular to the long axis of a rod of length ℓ and inertia $2m$ that is at rest on the surface. The ball collides with one end of the rod and sticks to it. (*a*) Just as the ball sticks, where is the center of mass of the ball-rod system? After the collision, what are (*b*) the magnitude and direction of the velocity of the center of mass and (*c*) the rotational speed of the system? Ignore any effects due to friction. ●●●
46. A 4.5-kg bowling ball is perched on a concrete ledge directly below your dorm room window, with the side of the ball opposite the holes touching the wall. Wanting to hold the ball in place so that it doesn't roll off and land on somebody, you manage to hook one of the holes with a wire and exert a purely tangential (and vertical) force on the ball. If the coefficient of static friction between ball and ledge is the same as that between ball and wall, $\mu_s = 0.50$, what is the maximum upward force you can exert so that the ball does not rotate and you lose your hold? Even though the ball has holes drilled in it, assume a uniform distribution of inertia. ●●●
47. Disk A in Figure P12.47 has radius R_A and thickness h and is initially rotating clockwise, as viewed from above, at $\omega_i/2$. Disk B, made of the same material as A, has radius $R_B = R_A/2$ and thickness h, and is initially rotating counterclockwise at ω_i. The two disks are constrained to rotate about the same axis, which runs through their centers. There is friction between the disks, and so once B slides down and touches A, they spin at the same rotational speed. (*a*) What is their common final rotational velocity? (*b*) If the frictional force is the only force that causes the system to lose energy, what is the fractional change in the kinetic energy? ●●●

Figure P12.47



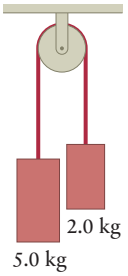
48. A 25-kg ladder of length 5.0 m leans against a smooth wall and makes an angle of 50° with the ground. A 75-kg man starts to climb the ladder. If the coefficient of static friction between ground and ladder is 0.50, what distance along the ladder can the man climb before the ladder starts to slip? ●●●

12.6 Rolling motion

49. An object rolls without slipping onto a surface where the coefficient of friction between object and surface is twice as great as that needed to prevent slipping. Describe the subsequent motion. ●
50. To get better acceleration, you want to reduce the inertia of your bicycle. You have your choice of spending $45 on only one of three light-alloy items, each reducing the inertia of your bicycle by the same amount: a new seat post, a new set of pedals, or a new set of wheel rims. Which is the best buy? ●
51. A 3.0-kg solid ball rolls without slipping down a ramp inclined to the horizontal at an angle of 30°. What are (*a*) the acceleration of the ball's center of mass and (*b*) the magnitude of the frictional force exerted on the ball? ●●
52. A 5.0-kg solid cylinder of radius 0.25 m is free to rotate about an axle that runs along the cylinder length and passes through its center. A thread wrapped around the cylinder is pulled straight from the cylinder so as to unwrap with a steady tensile force of 20 N. As the thread unwinds, the cylinder rotates and there is no slippage between thread and cylinder. If the cylinder starts from rest at $t_i = 0$, calculate (*a*) its rotational velocity at $t_f = 5.0$ s and (*b*) the angle through which it has turned at $t_f = 5.0$ s. Ignore any friction in the axle. ●●
53. A disk and a hoop roll down an inclined plane. If the plane is inclined at an angle of 30° from the horizontal, what is the minimum coefficient of friction required so that neither object slips? ●●
54. A 2.0-kg solid cylinder of radius 0.45 m rolls without slipping down a ramp inclined at an angle of 60° to the vertical. Calculate (*a*) the acceleration of the cylinder's center of mass, (*b*) the cylinder's rotational acceleration, and (*c*) the time interval required for it to travel 35 m down the incline if it starts from rest. (*d*) What is the rotational velocity of the cylinder at the end of the time interval found in part *c*? ●●
55. Two cans of pumpkin roll down a loading dock ramp. They have identical inertias, but one of them has a larger radius (and a shorter length) than the other. Which one, if either, makes it down the ramp first? ●●
56. You deliver a horizontal impulse to a cue ball with a cue stick while playing pool. If the ball is to roll without slipping, at what height h above its center (in terms of its radius R) must you strike it? ●●●
57. Two balls of the same radius and same inertia roll down an inclined plane, starting from rest. One ball is hollow, and the other is solid. What is the ratio of the time intervals the two balls require to reach the bottom? ●●

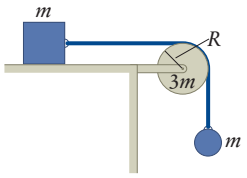
58. In the Atwood machine shown in Figure P12.58, the pulley radius is 0.10 m, and the rotational inertia of the pulley is 0.15 kg·m^2. Calculate (*a*) the acceleration of the blocks, (*b*) the tension in the cord on the left, and (*c*) the tension in the cord on the right. Ignore the cord's inertia. •••

Figure P12.58



59. A block of inertia m sits on a smooth surface (Figure P12.59). A light string is attached to it and placed over a solid pulley of inertia $3m$ and radius R, and a ball of inertia m is attached to the free end of the string. (*a*) If there is no slippage between string and pulley, determine the magnitude of the block's acceleration. (*b*) If you detach the ball and pull down on the hanging string with a force of magnitude mg, what is the magnitude of the block's acceleration? •••

Figure P12.59

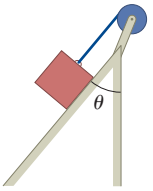


60. A 320-g can of soup is 10.8 cm tall and has a radius of 3.19 cm. (*a*) Calculate the can's rotational inertia about its axis of symmetry, assuming the can to be a solid cylinder. (*b*) When released from rest at the top of a ramp that is 3.00 m long and makes an angle of 25° with the horizontal, the can reaches the bottom in 1.40 s. What is the value of rotational inertia obtained from this experiment? (*c*) Compare the calculated and experimentally determined rotational inertias and suggest possible reasons for the difference. •••

12.7 Torque and energy

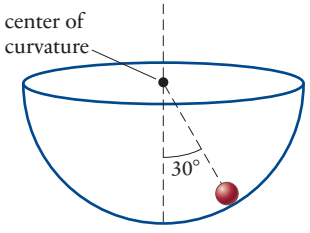
61. A 0.30-kg solid cylinder is released from rest at the top of a ramp 1.0 m long. The cylinder has a radius of 0.10 m, and the ramps is at an angle of 18° with the horizontal. What is the rotational kinetic energy of the cylinder when it reaches the bottom of the ramp? •
62. A solid 40 kg cylinder initially at rest has a radius of 0.10 m. What minimum work is required to get the cylinder rolling without slipping at a rotational speed of 25 s^{-1}? •
63. A 1.0-kg ring with an inner radius of 0.060 m and an outer radius of 0.080 m is sent rolling without slipping up a ramp that makes an angle of 30° with the horizontal. If the initial speed of the ring is 2.8 m/s, how far along the ramp does it travel before it comes to a stop? ••
64. A certain car engine delivers enough force to create 380 N·m of torque when the engine is operating at 3200 revolutions per minute. Calculate the average power delivered by the engine at this rotation rate. ••
65. A 680-kg disk of radius 1.2 m is mounted on a fixed axle. A force exerted on the disk causes a constant torque that gives the disk a rotational acceleration magnitude of 0.30 s^{-2} for 5.0 s. If the disk had an initial rotational speed of 4.5 s^{-1}, how much work was done to provide the torque, assuming no energy is lost to friction? ••
66. A large steel bar of length ℓ is hinged at one end to a wall. A mechanic holds the other end so that the bar is parallel to the ground and places a penny on the bar right at the end he is holding. (*a*) What is the rotational acceleration of the bar when he lets go? (*b*) Does the penny remain in contact with the bar after it is released? ••
67. A light, unstretchable string is wound around the perimeter of a 4.0-kg disk that has a radius of 0.50 m and is free to rotate about an axle that runs perpendicular to the disk face and through the disk center. A 2.0-kg block is connected to the free end of the string and allowed to slide down a ramp that makes an angle of $\theta = 37°$ with the vertical (Figure P12.67). Calculate the magnitude of the disk's rotational acceleration. Ignore friction between block and ramp. ••

Figure P12.67



68. A 5.0-kg hollow cylinder of radius 0.25 m rotates freely about an axle that runs through its center and along its long axis. A cord is wrapped around the cylinder and is pulled straight from the cylinder with a steady tensile force of 50 N. As the cord unwinds, the cylinder rotates, with no slippage between cord and cylinder. (*a*) Calculate the work done by the tensile force as the cylinder rotates through 1000 rad. (*b*) If the cylinder starts from rest, calculate its rotational speed after it has rotated through 1000 rad. ••
69. A marble of inertia m is held against the side of a hemispherical bowl as shown in Figure P12.69 and then released. It rolls without slipping. The initial position of the marble is such that an imaginary line drawn from it to the center of curvature of the bowl makes an angle of 30° with the vertical. The marble radius is $R_m = 10$ mm, and the radius of the bowl is $R_b = 100$ mm. Determine the rotational speed of the marble about its center of mass when it reaches the bottom. ••

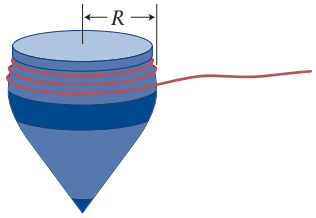
Figure P12.69



70. A 3.0-kg disk of radius 50 mm rolls down a ramp inclined at an angle of 28° with the vertical. If the disk starts out at rest and the coefficients of static and kinetic friction between the ramp and the disk are both 0.50, what is the rotational speed of the disk after it has traveled 1.5 m? ••

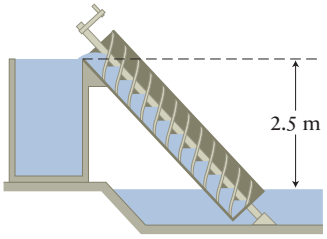
71. An almost-conical toy top (Figure P12.71) that has radius $R = 20$ mm and inertia 0.125 kg is spun up using a force of 5.0 N and a string that is 1.0 m long. What are (*a*) the work done on the top, (*b*) its kinetic energy, and (*c*) its final rotational speed? The shape factor for this top about its axis of symmetry is 0.35. ••

Figure P12.71



72. Archimedes' screw, one of the first mechanical devices invented for lifting water, is a very large screw fitted tightly inside a shaft (Figure P12.72). The bottom of the device is placed in a pool of water. As someone turns the handle to make the screw turn, water is carried up along the ridges of the screw and comes out the top of the shaft and into a storage tank. As the handle turns, the work done by the force exerted on the handle is converted to gravitational potential energy of the water-Earth system. Let's say you want to take a shower using this device. You figure that your shower will use 44 L of water, and so you have to raise this amount to the storage tank 2.5 m above the pool. Every time you turn the handle, you exert a force that causes a torque of 12 N·m. How many times must you turn the handle? ••

Figure P12.72



73. You attach a 0.50-m length of string to a 50-g puck and pass the other end of the string through a hole in the center of a table. Grasping the string under the table, you pull just enough string through the hole so that, when your friend gives the puck a sideways push, it moves in a circle of radius $r_i = 0.45$ m at 1.5 m/s. You then slowly pull on the string, decreasing the circle's radius to 0.20 m. (*a*) What is the puck's speed when $r = 0.20$ m? (*b*) What is the tension in the string as a function of r? What is its value when $r = 0.20$ m? (*c*) How much work is done in moving the puck from the 0.45-m circle to the 0.20-m circle? ••

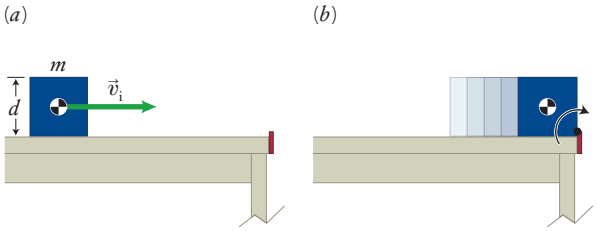
74. A 30-kg solid sphere of radius 0.12 m is rolling without slipping on a horizontal surface at 2.0 m/s. (*a*) What average torque is required to stop the sphere in 5.0 rev without inducing skidding? (*b*) If this torque is caused by a soft braking bumper that is lowered down until it just makes contact with the top of the sphere, what is the magnitude of the frictional force between bumper and sphere? •••

75. A marble that has a radius of 10 mm is placed at the top of a globe of radius 0.80 m. When released, the marble rolls without slipping down the globe. Determine the angle from the top of the globe to the location where the marble flies off the globe. •••

76. A 4.0-kg bowling ball is thrown down the alley with a speed of 10.0 m/s. At first the ball slides with no rotation. The coefficient of friction between ball and alley surface is 0.20. (*a*) How long a time interval does it take for the ball to achieve pure rolling motion? (*b*) What is the ball's translational speed at the end of this time interval? •••

77. You shove a cube of inertia m and side length d so that it slides along a smooth table with speed v_i (Figure P12.77*a*). The cube then hits a raised lip at the end of the table. After it hits the lip, the cube begins to rotate about it (Figure 12.77*b*). (*a*) Show that the magnitude of the cube's angular momentum about the lip before the collision is $L = mdv_i/2$. (*b*) Explain why the angular momentum still has that value at the instant of collision, before the cube has had time to rotate much. (*c*) What is the rotational acceleration of the cube the instant after it hits the lip? (*d*) What maximum initial speed can the cube have so that it does not topple over the lip? •••

Figure P12.77



12.8 The vector product

78. Which of the following expressions make sense: (*a*) $\vec{A} \cdot (\vec{B} \times \vec{C})$, (*b*) $\vec{A} \times (\vec{B} \cdot \vec{C})$, (*c*) $\vec{A} \times (\vec{B} \times \vec{C})$? •

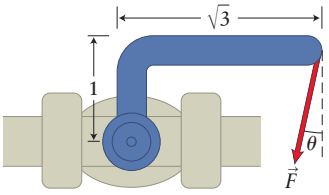
79. When you want to turn a bicycle left, the first thing you do is lean left. Why is it important to do this? •

80. If the magnitude of the vector product of two vectors is the same as the magnitude of the scalar product of the same two vectors, what is the angle between them? •

81. A cyclist exerts a vertical force on a bike pedal. This pedal is at the end of a crank that is 0.20 m long and pivoted to rotate about the axle of the chain wheel. If the cyclist pushes downward with a force of 150 N, determine the magnitude of the torque caused by this force when the crank is (*a*) horizontal, (*b*) 30° from the horizontal, (*c*) 45° from the horizontal, (*d*) 60° from the horizontal, and (*e*) vertical. (*f*) Calculate the average magnitude of this torque exerted by one of the cyclist's feet during one cycle of the pedals, assuming there is no clip holding the cyclist's feet to the pedals. ••

82. The two arms of the L-shaped handle on the spigot of Figure P12.82 have a length ratio of $1:\sqrt{3}$. At what angle θ do you want to pull down on the end of the handle to maximize the torque your force causes? ●●

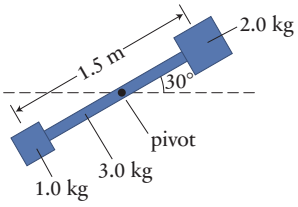
Figure P12.82



83. Show that, if vectors $\vec{A}$ and $\vec{B}$ are both in an xy plane, $\vec{A} \times \vec{B}$ is a vector that is perpendicular to the xy plane and has magnitude $|\vec{A} \times \vec{B}| = |A_x B_y - A_y B_x|$. ●●

84. A 3.0-kg rod that is 1.5 m long is free to rotate in a vertical plane about an axle that runs through the rod's center, is perpendicular to the rod's length, and runs parallel to the floor. A 1.0-kg block is attached to one end of the rod, and a 2.0-kg block is attached to the other end. At some instant, the rod makes an angle of 30° with the horizontal so that the blocks are in the positions shown in Figure P12.84. (*a*) Determine the torque caused by the forces exerted on the system at this instant. (*b*) Determine the rotational acceleration of the system at this instant. Ignore friction and assume the blocks are small enough that any length they add to the rod can be ignored. ●●

Figure P12.84

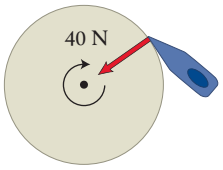


85. A motor drives a disk initially at rest through 23.9 rotations in 5.0 s. Assume the vector sum of the torques caused by the force exerted by the motor and the force of friction is constant. The rotational inertia of the disk is $4.0\ \text{kg} \cdot \text{m}^2$. When the motor is switched off, the disk comes to rest in 12 s. (*a*) What is the torque created by the force of friction? (*b*) What is the torque caused by the force exerted by the motor? ●●

86. The angular momentum of the propellers of a small airplane points directly forward from the plane. (*a*) In what direction do the propellers rotate as seen from the rear of the plane? (*b*) If the plane is flying horizontally and suddenly pulls upward, in which direction does the nose of the plane tend to move? Justify your answer. ●●

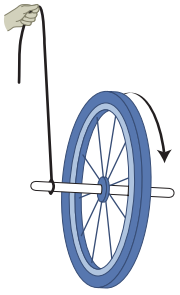
87. A 40-kg sharpening wheel of radius 0.10 m is rotating at 3.3 revolutions per second. A 6.0-kg axe is pressed against the rim with a force of 40 N directed as shown in Figure P12.87. Treat the wheel as a disk, and assume a coefficient of kinetic friction between wheel and axe blade of $\mu_k = 0.35$. (*a*) If there is no power source keeping the wheel rotating, how long does it take to stop? (*b*) How many revolutions does the wheel make while it is slowing down? ●●

Figure P12.87



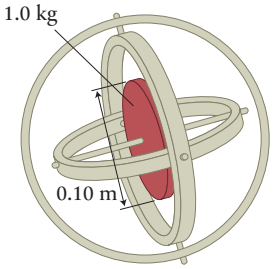
88. A heavy-rimmed bicycle wheel is set spinning in the direction shown in Figure P12.88. A string is tied to one end of the axle, and someone is holding up the string. (*a*) Use torque arguments to explain why the wheel slowly revolves horizontally around the string end of the axle when the free end of the axle is released. (This effect is called *precession*.) (*b*) The precession speed increases as time passes. Explain why. ●●●

Figure P12.88



89. A gyroscope consists of a spinning wheel mounted inside a support structure (Figure P12.89). Because of conservation of angular momentum, the gyroscope keeps its orientation in space unless some outside force causes a torque on it. Suppose the 1.0-kg spinning wheel of a certain gyroscope has a diameter of 0.10 m and is spinning at 14,200 revolutions per minute. What is the magnitude of the torque required to tilt the gyroscope 90° in 20 s such that the spin speed is still 14,200 revolutions per minute when the gyroscope is in its new orientation? ●●●

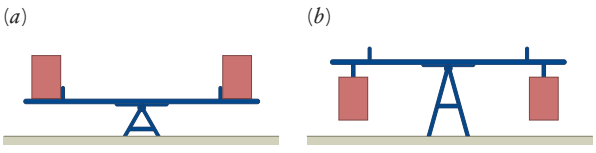
Figure P12.89



Additional Problems

90. Two identical boxes are placed on opposite ends of the board of a playground seesaw as in Figure P12.90*a*, so that the system is in mechanical equilibrium. (*a*) What happens to that equilibrium if you exert a small, brief force, either upward or downward, on either end of the board? (*b*) With the boxes attached to the underside of the board as in Figure P12.90*b* and the system again in mechanical equilibrium, what happens to the equilibrium if you exert such a force on either end? ●

Figure P12.90



91. For the x and y axes of a right-handed Cartesian coordinate system, what is the direction of $\hat{\imath} \times \{\hat{\imath} \times [\hat{\imath} \times (\hat{\imath} \times \hat{\jmath})]\}$? •

92. Standing on a round raft floating on a pond, how do you turn the raft around 180°? •

93. Dragster drivers have to avoid supplying too much power to the vehicle because too much power causes the front end to rise in a "wheelie," compromising steering control. (*a*) Why does this happen? (*b*) What advantage comes from having front-wheel drive in a dragster? ••

94. If everyone on Earth simultaneously walked from west to east, by what fraction would the length of the day change? Would it lengthen or shorten? ••

95. A 51-kg box is suspended from the right end of a horizontal rod that has very small inertia. The left end of the rod is affixed to a wall by a pin. A wire connects the right end of the rod to the wall directly above the pin, making an angle of 40° with the rod. (*a*) Calculate the tension in the wire. (*b*) Determine the magnitude of the reaction force the pivot exerts on the rod. (*c*) Repeat parts *a* and *b* for a 10.2-kg rod. ••

96. A 35-kg child stands on the edge of a playground merry-go-round that has a radius of 2.0 m and a rotational inertia of 500 kg · m^2. The merry-go-round has a rotational speed of 0.20 s^{-1} when the child is standing still. If the rotational speed increases to 0.25 s^{-1} when the child starts walking along the edge, with what speed is he walking in the reference frame of the merry-go-round, and in which direction? ••

97. A neutron star is the compact remnant of a very large star that has exploded. Suppose that right after one such explosion, when much of the star's inertia has been blasted away, the star's remaining core has an inertia of 4×10^{30} kg, a radius of 13×10^8 m, and completes one revolution every five days. What is the new spin rate once the core gravitationally collapses into a neutron star of radius 20 km? Assume negligible loss of inertia during the collapse. ••

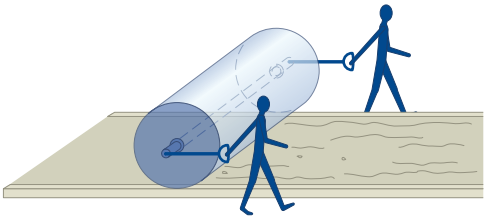
98. A horizontal 2.0-kg rod is 2.0 m long. An 8.0-kg block is suspended from its left end, and a 4.0-kg block is suspended from its right end. (*a*) Determine the magnitude and direction of the single extra force necessary to keep the rod in mechanical equilibrium. (*b*) At what distance from the left end of the rod must the point of application of this force be? ••

99. What maximum torque can a bicyclist deliver to the pedals? ••

100. A solid ball of inertia m rolls without slipping down a ramp that makes an angle θ with the horizontal. (*a*) What frictional force is exerted on the ball? (*b*) As a function of θ, what coefficient of friction is required to prevent slipping? ••

101. You make a lawn roller out of a 125-kg solid cylinder, set to rotate about a central axle. You rig up handles on either side of the axle so that you and a friend can pull on the handles horizontally to pull the roller in a direction perpendicular to its rotation axis (Figure P12.101). (*a*) If each of you pulls with a horizontal force of 500 N, what is the magnitude a of the roller's acceleration? (*b*) What is the minimum coefficient of static friction necessary to keep the roller from slipping? ••

Figure P12.101



102. A marble is shot across a smooth wooden floor so that the marble both rotates about a horizontal axis and slides. If at a certain instant the marble's rotational kinetic energy equals the translational kinetic energy of its center of mass, what is the ratio of its center-of-mass speed to the speed of a point on its "equator" associated with rotation? ••

103. A 12-kg cylinder of radius 0.10 m starts at rest and rolls without slipping down a ramp that is 6.0 m long and inclined at 30° to the horizontal. When the cylinder leaves the end of the ramp, it drops 5.0 m to the ground. At what horizontal distance from the end of the ramp does it land? ••

104. A child runs in a straight line tangent to a playground carousel, jumps on, and holds tight as the carousel begins to rotate. Which of the following statements is correct for this collision? (*a*) The coefficient of restitution is greater than 1 because there is now rotational energy where there was none before. (*b*) The coefficient of restitution is 1 because the system's translational kinetic energy is completely converted to rotational kinetic energy. (*c*) The coefficient of restitution is less than 1 because the sum of the translational and rotational kinetic energies in the final state is less than the sum in the initial state. (*d*) The coefficient of restitution cannot be defined in a system in which there is both translational motion and rotational motion. (*e*) Explain why all the answers except the one you chose are wrong. ••

105. In a judo hip throw, you pull your 60-kg opponent onto your back to bring her center of mass just above your hip and then rotate her about your hip. Assume that the lever arm distance of your grip on her is 0.30 m from your hip and that her rotational inertia about your hip is 15 kg · m^2. (*a*) What is the magnitude of the pulling force you must exert on her to cause a rotational acceleration of 1.0 s^{-2}? (*b*) What is the magnitude of the pulling force if you throw her from a more upright position, so that the gravitational force has a lever arm of 0.12 m about your hip? •••

106. You've been asked to design a flywheel and the associated emergency friction braking system for an electric vehicle. In order to fit the available space, the flywheel must be cast in the form of a solid steel disk with a thickness of 50 mm and a radius that cannot exceed 0.20 m. For safety purposes, the emergency braking system must slow the flywheel from 6000 revolutions per minute to 0 in 12 revolutions. You know that typical braking systems utilize a pair of "shoes" mounted on calipers to squeeze the thickness of the disc symmetrically near its edge, but you are concerned about how much force the brake shoes might have to exert on the disk. ••• CR

107. A 3.0-kg block is attached to one end of a light, unstretchable string that is wrapped securely around a cylinder that has a 0.30-m radius and a rotational inertia of 0.80 $kg \cdot m^2$. The cylinder is free to rotate about an axle aligned along its long axis. The block is held right alongside the cylinder (so there is essentially no unwound string) and then released from rest. As the block falls, it unwinds the string from the cylinder, with no slippage between string and cylinder. (*a*) Calculate the magnitude of the block's acceleration. (*b*) Calculate the rotational speed of the cylinder after the block has dropped 1.5 m. (Hint: First use the acceleration from part *a* to determine the block's speed.) (*c*) Use energy methods to obtain the cylinder's rotational speed after the block has dropped 1.5 m. (*d*) Calculate the tension in the string after the block has dropped this distance. (*e*) Calculate the instantaneous power delivered by the string and block to the cylinder after the block has dropped this distance. ●●●

108. You are building a pool table and want the design to be such that any ball that rolls into one of the side rails without slipping bounces off the rail and rolls away in another direction, again without slipping. After some experimenting, you decide that you must build the rails to a specific height that is some function of the radius of the balls. (Hint: It seems plausible to assume that the collision is elastic.) ●●● CR

109. A yo-yo is composed of two disks of radius a, with many layers of string wrapped around the axle. The radial distance from the center of the axle to the top layer of string at any instant is b. With the end of the string wrapped around your finger and $b/a \approx 1$, you let the yo-yo drop, unwinding string and ultimately spinning pretty fast when it reaches the bottom of the string $(b/a \approx 0)$. You begin to wonder how the motion of the yo-yo as the string unwinds depends on the value of b. ●●● CR



13

Gravity

13.1 Universal gravity
13.2 Gravity and angular momentum
13.3 Weight
13.4 Principle of equivalence

13.5 Gravitational constant
13.6 Gravitational potential energy
13.7 Celestial mechanics
13.8 Gravitational force exerted by a sphere

CONCEPTS

QUANTITATIVE TOOLS

The motion of celestial bodies has fascinated humankind since antiquity. What keeps Earth's Moon, the planets, and other heavenly objects moving along their orbits? Motivated largely by aesthetics, Greek philosophers proposed that celestial bodies move in circles because they believed that circular motion was the most "natural" and the most perfect form of motion. As we have seen, however, moving objects travel in a straight line when they are not subject to any external forces. In the early 17th century, the German astronomer Johannes Kepler suggested that the Sun was somehow driving the planets around their orbits.

If the laws of physics are truly universal, then phenomena on a cosmic scale should follow the same laws that govern objects around us. In this chapter we see that we can indeed apply our notions of mechanics to describe the motion of celestial bodies.

13.1 Universal gravity

Careful analysis of the motion of the Moon and the planets shows that the Moon revolves around Earth and that Earth and the other planets revolve around the Sun at roughly constant speed in nearly circular orbits.* In other words, to a good approximation they move in circular motion at constant speed. These bodies generally rotate about an axis through their centers in addition to revolving around the Sun, but in this chapter we are concerned with their *orbital* motion—that is, rotational quantities associated with the revolving motion of objects.

As we saw in Chapter 11, any object that is in circular motion at constant speed has a centripetal acceleration. As illustrated in **Figure 13.1**, in the reference frame of the Sun, Earth has an acceleration that points toward the Sun; in the reference frame of Earth, the Moon has an acceleration that points toward Earth.

To get some feel for the quantities involved in celestial mechanics, do the following checkpoint before proceeding.

13.1 The orbital period of the Moon around Earth is 27.32 days; that of Earth around the Sun is 365.26 days. (*a*) Which orbit has the greater rotational speed: Earth's or the Moon's? Are the orbital rotational speeds of Earth and the Moon around the Sun great or small relative to the rotational speeds you encounter in daily life? (*b*) The radius of the Moon's orbit around Earth is $R_M = 3.84 \times 10^8$ m; that of Earth's orbit around the Sun is $R_E = 1.50 \times 10^{11}$ m. Which has the greater speed? (*c*) Calculate the centripetal acceleration of Earth in the reference frame of the Sun and the centripetal acceleration of the Moon in the reference frame of Earth. (*d*) How do these accelerations compare with the acceleration due to gravity near Earth's surface?

As we saw in Chapter 11, some force is required to supply a centripetal acceleration. All examples we have considered so far involve contact forces. Neither the Moon nor Earth is in contact with anything, however, and so the force that holds them in orbit must be a field force. In the late 17th century, Isaac Newton made the bold proposal that the force that holds celestial bodies in orbit is a gravitational force—the same force that causes objects near Earth's surface to fall (**Figure 13.2**). Every piece of matter in the universe—feather or planet—attracts every other piece of matter. In other words, a gravitational attraction between Earth and the Moon is responsible for the centripetal acceleration of the Moon, just as a gravitational attraction between a terrestrial object and Earth is responsible for the free

Figure 13.1 Both Earth and the Moon have nearly circular orbits and thus have a centripetal acceleration.

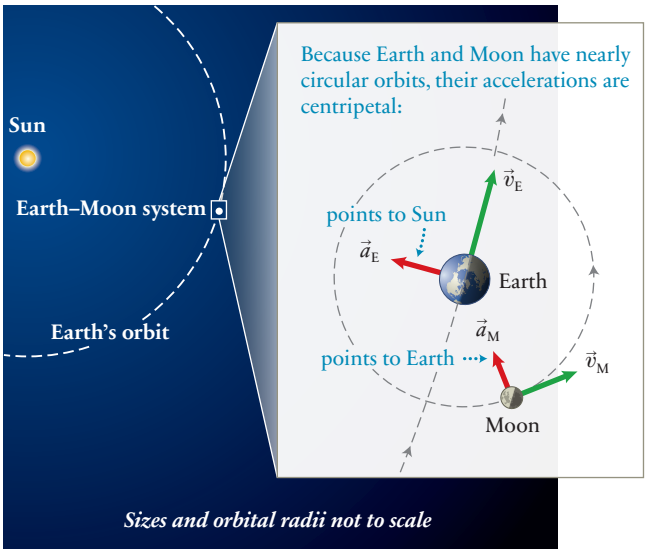
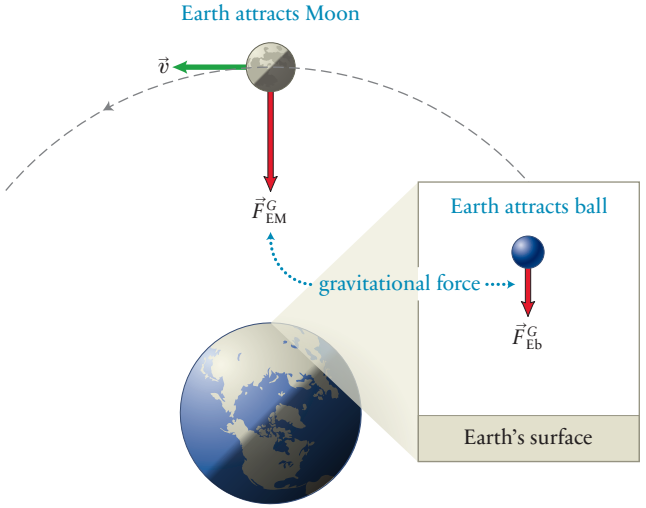


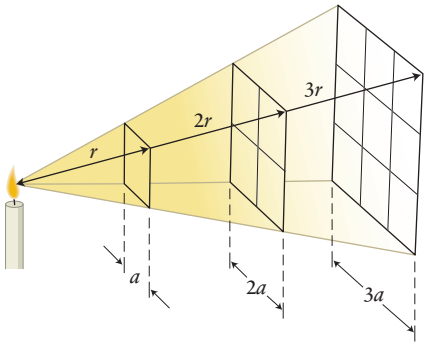
Figure 13.2 The invisible tether that holds the Moon in orbit is, in reality, the same kind of gravitational interaction that causes a ball to fall near Earth's surface. The Moon's orbit is not to scale.



*Earth, for example, moves along an orbit that, to the eye, is indistinguishable from a circle. The speed of Earth along its orbit is constant to within 3%.

CONCEPTS

Figure 13.3 Light that strikes a surface area a^2 located a distance r from the source strikes a surface of area $2a \times 2a = 4a^2$ at a distance $2r$ and a surface of area $3a \times 3a = 9a^2$ at a distance $3r$. The intensity—the amount of light per unit area—decreases with distance r as $1/r^2$.



fall of that object. The reason the Moon does not fall to Earth like a freely falling ball is that the Moon has just the right speed to keep moving in a circle around Earth.

Newton furthermore postulated that the effect of gravity weakens with distance, much as the effect of a magnet weakens with distance. If gravity spreads out uniformly from an object, as the rays from a light source do, then the effect of gravity should decrease as $1/r^2$—the inverse of the square of the distance (**Figure 13.3**). The radius of the Moon's orbit is 60 times the radius of Earth, so the gravitational force exerted by Earth on the Moon should be smaller by a factor of $\left(\frac{1}{60}\right)^2$ than the gravitational force exerted by Earth on an object near the ground. An object at the same distance from Earth as the Moon should therefore experience an acceleration due to gravity of magnitude $\frac{1}{3600}g = (9.8\ \text{m/s}^2)/3600 = 0.00272\ \text{m/s}^2$. This value is in excellent agreement with the Moon's centripetal acceleration (see Checkpoint 13.1), confirming the $1/r^2$ dependence of the gravitational force.

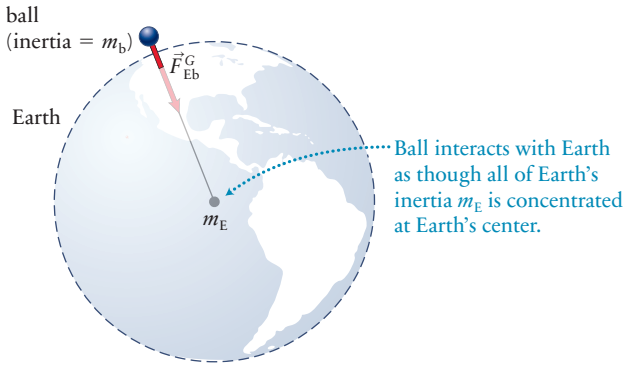
There is one catch in this reasoning: The dilution of the gravitational force by a factor of $1/r^2$ assumes that the object near Earth's surface is one Earth radius away from the point of attraction. In other words, a ball falling near the surface of Earth interacts with Earth as if all the matter were concentrated at its center—6378 km below the surface—as illustrated in **Figure 13.4*a***. Yet if all matter interacts gravitationally with a $1/r^2$ dependence, then a piece of Earth-matter near the surface of Earth attracts the ball much more strongly than does a piece of Earth-matter deep inside or on the opposite side of Earth (Figure 13.4*b*). It is certainly not obvious that the forces exerted by all the different pieces of Earth add up to give the same result as in Figure 13.4*a*. Newton, however, proved mathematically that:

A uniform solid sphere exerts a gravitational force outside the sphere with a $1/r^2$ dependence as if all the matter in the sphere were concentrated at its center.

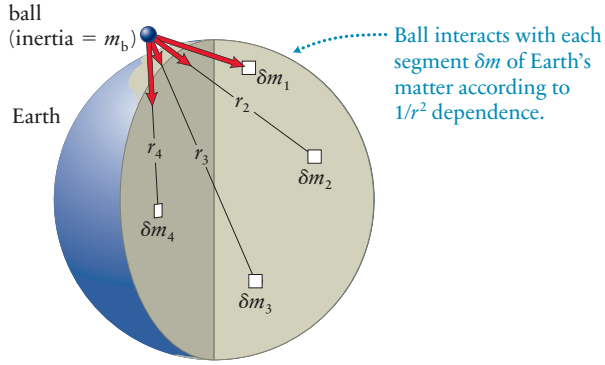
The proof of this statement is given in Section 13.8. A direct consequence of this statement is that the distance between gravitationally interacting spherical objects should always be taken to be the distance between their centers.

Figure 13.4 (*a*) A solid sphere exerts a gravitational force as if all of the sphere's matter were concentrated at the center of the sphere. (*b*) Because the magnitude of the gravitational force between two objects depends on the distance between them, the gravitational forces exerted by the various "chunks" of Earth's interior represented by δm have different magnitudes.

(*a*) Gravitational interaction between Earth and ball near Earth's surface



(*b*) Reason gravitation acts as though from Earth's center



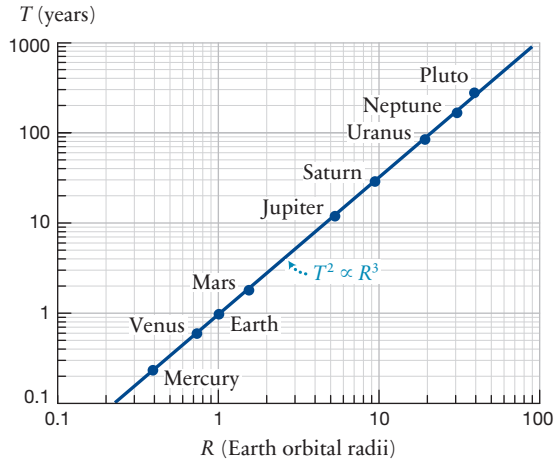
13.2 If the force of gravity decreases with the inverse square of the distance, why were we allowed, in all our earlier work on the gravitational force, to say that an object sitting on the ground, an object sitting in a tree 10 m above the ground, and an object flying at an altitude of 10 km all experience the same 9.8-m/s^2 acceleration due to gravity?

Further evidence of the $1/r^2$ dependence of the gravitational force is provided by the observed relationship between the radii R and the periods T of the planetary orbits. As it orbits the Sun, each planet has centripetal acceleration $a_c = v^2/R$, where $v = 2\pi R/T$ is the planet's linear speed and R is the radius of its orbit. So $a_c = 4\pi^2 R/T^2$ or $a_c \propto R/T^2$.* If this centripetal acceleration is provided by a gravitational attraction exerted by the Sun on the planet, the magnitude of the force (and hence the centripetal acceleration) should decrease

*The symbol $\propto$ means "is proportional to." That is, $a_c \propto R/T^2$ means that there is some constant k such that $a_c = kR/T^2$.

CONCEPTS

Figure 13.5 The planetary period T versus the orbital radius R. The data are plotted on a logarithmic grid to fit the large range of values. To convince yourself of the $T^2 \propto R^3$ dependence, compare the values of T at $R = 1$ Earth orbital radius and $R = 100$ Earth orbital radii.



as $1/R^2$. So, the $1/r^2$ dependence of the gravitational force requires that $a_c \propto 1/R^2$. Combining this result with $a_c \propto R/T^2$, we get $1/R^2 \propto R/T^2$, or $T^2 \propto R^3$. In words:

The square of the period of a planetary orbit is proportional to the cube of the orbit's radius.

Johannes Kepler had established this fact from astronomical data in 1619, providing support for the $1/r^2$ dependence of the gravitational force proposed later by Newton. The excellent agreement with astronomical observations is shown in **Figure 13.5**.

13.3 Suppose the universe were two dimensional rather than three dimensional. (*a*) Following the line of reasoning illustrated in Figure 13.3, describe how the strength of the gravitational force would depend on distance in this flat universe. (*b*) How would the periods of the planetary orbits be related to their radii in this universe?

CONCEPTS

What determines the strength of the gravitational attraction? To answer this question, consider two identical lumps of clay, each of inertia m. If we let the lumps fall freely, each is subject to a single downward force of gravity and undergoes an acceleration g. To give each lump this acceleration, Earth must exert a force mg on it. Now imagine combining the two lumps. The double lump has twice the inertia, which means it is twice as hard to accelerate as either single lump. Yet, like any other object near Earth's surface, the double lump undergoes an acceleration g. From this observation, we conclude that Earth pulls twice as hard on the double lump as it does on either single lump. Apparently the gravitational pull on an object is proportional to the quantity of material in it, a quantity called the **mass** of the object. The double lump has twice the mass of either single lump.

The gravitational pull exerted by Earth on an object is proportional to the object's mass.

The acceleration of a freely falling object near Earth's surface does not depend on the type of material: In the absence of air resistance, all objects, regardless of their composition, configuration, or any other property of the object, have the same acceleration in free fall. This observation suggests that the inertia of an object and the force of gravity exerted on it are proportional to each other. Given that the force of gravity is proportional to the object's mass, we can consider mass and inertia to be equal and let the SI unit of mass be the kilogram (kg). As we shall see in the next chapter, this equality between mass and inertia breaks down for motion at very high speeds. In everyday situations, however, we can say:

The mass of an object is equal to the object's inertia.

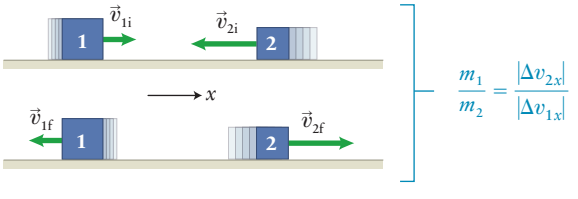
For the description of everyday situations, therefore, we can denote both by the same symbol m and express each in units of kilograms. Unless we need to refer explicitly to an object's resistance to a change in velocity or account for effects that occur at very high speeds, we henceforth shall call both simply the *mass* of the object.

The equality of inertia and mass may seem obvious, but consider the two experiments shown in **Figure 13.6**. The collision experiment in Figure 13.6*a* allows us to determine the ratio of the objects' inertias. Now suppose we suspend the same two objects, from a beam that is supported at its center of mass. We arrange the objects' suspension points so that the beam is balanced as shown in Figure 13.6*b*. Experimentally we discover that for *any* two objects, the ratio m_1/m_2 determined in the first experiment is exactly equal to the ratio r_2/r_1 determined in the second experiment. As we saw in Chapter 12, when the beam is balanced, the torques on the beam add up to zero, so $r_1 F_{E1}^G = r_2 F_{E2}^G$, which means that

$$\frac{F_{E1}^G}{F_{E2}^G} = \frac{r_2}{r_1}.$$

Figure 13.6 Establishing the equivalence between inertia and mass.

(*a*) Collision between objects gives ratio of inertias



(*b*) Beam in equilibrium gives ratio of gravitational forces

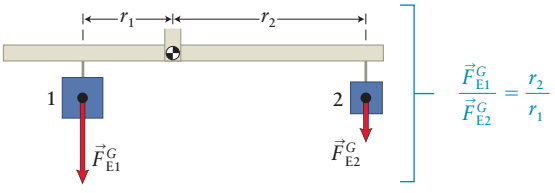
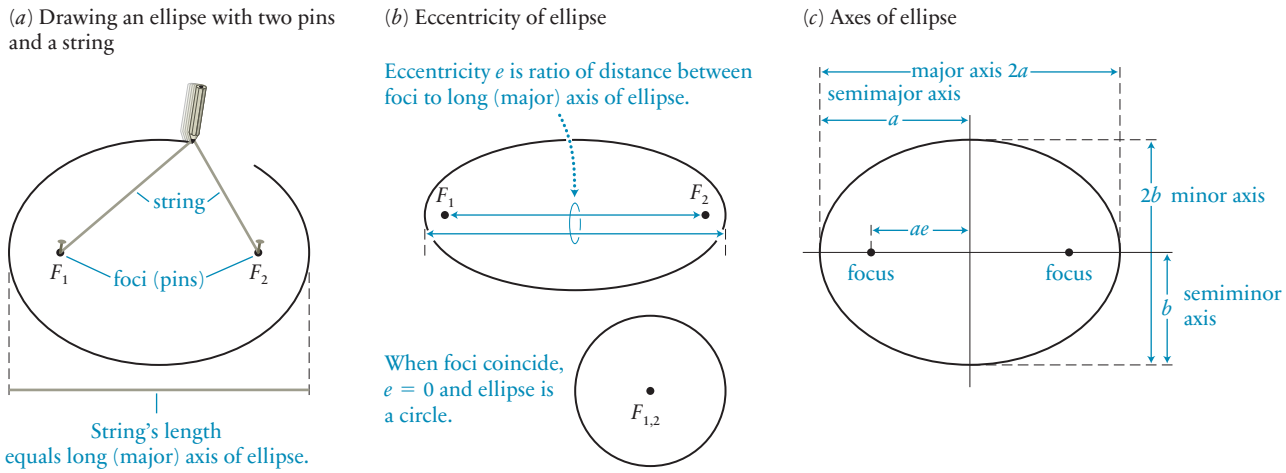


Figure 13.7 Tracing an ellipse. The greater the separation of the foci, the more eccentric (elongated) the ellipse ($0 < e < 1$). In the special case where the two foci coincide ($e = 0$), the ellipse becomes a circle of radius a.



Combining this with the experimental observation that

$$\frac{m_1}{m_2} = \frac{r_2}{r_1},$$

we see that

$$\frac{F^G_{E1}}{F^G_{E2}} = \frac{m_1}{m_2}.$$

In words, for any two objects, the ratio of the gravitational forces they experience is equal to the ratio of their inertias. If you think about it, there is no reason to expect this result. The two experiments are completely different: The collision measures the objects' tendencies to resist a change in velocity—their *inertias*—and involves no gravity; the balancing experiment measures the gravitational pull on the objects—their *mass*—and involves no motion. This equivalence is therefore truly remarkable.

Because the gravitational force exerted by Earth on an object is proportional to that object's mass, we can expect the gravitational force exerted by Earth on the Moon to be proportional to the Moon's mass: $F^G_{EM} \propto m_M$. The attraction between Earth and the Moon is mutual, and so symmetry requires the force of gravity exerted by the Moon on Earth to be proportional to the mass of Earth: $F^G_{ME} \propto m_E$. In addition, we know that $F^G_{EM} = F^G_{ME}$, and so the force of gravity must be proportional to the product $m_M m_E$.

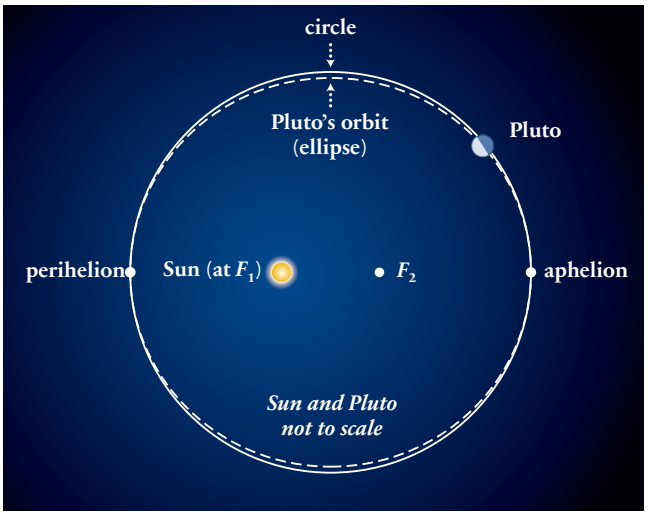
Combining this conclusion with the $1/r^2$ dependence, we can now write for the gravitational attraction between any two material objects of masses m_1 and m_2 separated by a distance r:

$$F^G_{12} \propto \frac{m_1 m_2}{r^2}.$$

This is Newton's **law of gravity.** It applies to all the mass in the universe: Every piece of matter attracts every other piece of matter in the universe according to this law.

The $1/r^2$ dependence also explains the shape of the planetary orbits. Using conservation of energy and angular momentum, we can show that the orbit of a body moving under the influence of gravity must be an ellipse (**Figure 13.7**), a circle, a parabola, or a hyperbola. Most comets, for example, move in highly elliptical orbits with the Sun at one focus of the ellipse. The planets also move in elliptical orbits with the Sun at one focus. As you can see from the data in **Table 13.1**, however, the planetary orbits are close to circular, even for dwarf planet Pluto, whose orbit has the greatest eccentricity of all. Even though the effect of the eccentricity on the shape of the orbit is minimal, its effect on the speed is considerable. The Sun is appreciably off-center for Pluto's orbit, and so the distance to the Sun and the speed vary as the planet moves between the closest point to the Sun, the *perihelion*, and the farthest point, the *aphelion* (**Figure 13.8**).

Figure 13.8 In spite of having the greatest eccentricity in Table 13.1, Pluto's orbit (dashed line) is remarkably close to a circle (solid line). Because the Sun is off-center, Pluto's speed varies appreciably along the orbit.



CONCEPTS

Table 13.1 Solar system data (in SI units and relative to Earth)

	Mass		Equatorial radius		Orbit†				
					semimajor axis		eccentricity	period	
	(kg)	(m_E)	(m)	(R_E)	(m)	(a_E)		(s)	(years)
Sun	2.0×10^{30}	3.3×10^5	7×10^8	110	–	–	–	–	–
Mercury	3.30×10^{23}	0.06	2.440×10^6	0.38	5.79×10^{10}	0.39	0.206	7.60×10^6	0.24
Venus	4.87×10^{24}	0.81	6.052×10^6	0.95	1.082×10^{11}	0.72	0.007	1.94×10^7	0.62
Earth	5.97×10^{24}	1	6.378×10^6	1	1.496×10^{11}	1	0.017	3.16×10^7	1
Mars	6.42×10^{23}	0.11	3.396×10^6	0.53	2.279×10^{11}	1.52	0.09	5.94×10^7	1.88
Jupiter	1.90×10^{27}	318	7.149×10^7	11.2	7.783×10^{11}	5.20	0.05	3.74×10^8	11.86
Saturn	5.68×10^{26}	95.2	6.027×10^7	9.45	1.427×10^{12}	9.54	0.05	9.29×10^8	29.45
Uranus	8.68×10^{25}	14.5	2.556×10^7	4.01	2.871×10^{12}	19.2	0.05	2.65×10^9	84.02
Neptune	1.02×10^{26}	17.1	2.476×10^7	3.88	4.498×10^{12}	30.1	0.01	5.20×10^9	164.8
Pluto	1.31×10^{22}	0.002	1.151×10^6	0.18	5.906×10^{12}	39.5	0.25	7.82×10^9	247.9
Moon	7.3×10^{22}	0.012	1.737×10^6	0.27	3.84×10^8	0.0026	0.055	2.36×10^6	0.075

†The elliptical orbits of the planets and the Moon are specified by their *semimajor axis a* (half the major axis) and eccentricity *e*; see Figure 13.7. With the exception of Mercury and Pluto, the eccentricity is small and so the orbits are close to being circular.

13.4 Suppose the force of gravity between two objects 1 and 2 of masses m_1 and m_2 were proportional to the sum $m_1 + m_2$ rather than to the product $m_1 m_2$. (*a*) Would this dependence be consistent with the following two requirements? (*i*) The force exerted by 1 on 2 is equal in magnitude to the force exerted by 2 on 1. (*ii*) The force of gravity exerted by Earth on two identical objects is equal to the sum of the forces exerted on the individual objects. (*b*) Can you think of any other combination of m_1 and m_2 that satisfies requirements (*i*) and (*ii*)? (*c*) Suppose you are given a lump of matter of mass *M*. How should you divide this lump into two parts to maximize the force of gravity between the two parts?

CONCEPTS

EXAMPLE 13.1 Comparing gravitational pulls

Compare the gravitational force exerted by Earth on you with (*a*) that exerted by a person standing 1 m away from you and (*b*) that exerted by Earth on Pluto.

1 GETTING STARTED Although I don't know yet how to calculate a numerical value for the gravitational force, I know that the gravitational force between two objects with masses m_1 and m_2 separated by a distance r is proportional to the factor $m_1 m_2 / r^2$. To compare gravitational forces, I should therefore compare these factors for each given situation.

2 DEVISE PLAN My mass is about 70 kg, and I can get the masses of Earth and Pluto from Table 13.1. When I stand on the surface of Earth, the distance between me and the center of the planet is its radius, which is given in Table 13.1. To do part *b*, I need to know the distance between Pluto and Earth, which is not given in the table. Planetary orbits are very nearly circular, however, so I can consider the semimajor axis *a* to be the radius of each (nearly) circular orbit. Table 13.1 shows that the "radius" of Pluto's orbit is about 40 times greater than that of Earth's orbit, and so I make only a small error by taking the Sun-Pluto distance as a measure of the (average) Earth-Pluto distance. Armed with this information, I'll calculate the factor $m_1 m_2 / r^2$ first for Earth and me, then for the two situations described in the two parts of this problem.

3 EXECUTE PLAN When I stand on the surface of Earth, $r = R_E$, and so the factor $m_1 m_2 / r^2$ is

$$\frac{m_1 m_E}{R_E^2} = \frac{(70\ \text{kg})(5.97 \times 10^{24}\ \text{kg})}{(6.38 \times 10^6\ \text{m})^2} = 1.0 \times 10^{13}\ \text{kg}^2/\text{m}^2.$$

(*a*) For two 70-kg people separated by 1 m, I get

$$\frac{m_1 m_2}{r_{12}^2} = \frac{(70\ \text{kg})^2}{(1\ \text{m})^2} = 4.9 \times 10^3\ \text{kg}^2/\text{m}^2.$$

Thus the gravitational force exerted by a person standing 1 m from me is $(4.9 \times 10^3\ \text{kg}^2/\text{m}^2)/(1.0 \times 10^{13}\ \text{kg}^2/\text{m}^2) = 4.9 \times 10^{-10}$ times the gravitational force exerted by Earth on me. ✔

(*b*) For Earth and Pluto, I have

$$\frac{m_E m_P}{r_{EP}^2} = \frac{(5.97 \times 10^{24}\ \text{kg})(1.36 \times 10^{22}\ \text{kg})}{(5.9 \times 10^{12}\ \text{m})^2} = 2.3 \times 10^{21}\ \text{kg}^2/\text{m}^2.$$

This is $(2.3 \times 10^{21}\ \text{kg}^2/\text{m}^2)/(1.0 \times 10^{13}\ \text{kg}^2/\text{m}^2) = 2.3 \times 10^8$ times greater than the attraction between Earth and me. ✔

4 EVALUATE RESULT That the gravitational force exerted by another person on me is more than a billion times smaller than that exerted by Earth on me makes sense: Only the gravitational attraction between Earth and objects on Earth is noticeable. That the gravitational force exerted by Earth on Pluto is 200,000,000 times greater than that exerted by Earth on me, even though Pluto is 1,000,000 times farther from Earth's center than I am, is amazing. Of course, Pluto's mass is about 10^{20} times greater than mine, and that factor more than makes up for the large difference between the Earth-me distance and the Earth-Pluto distance.

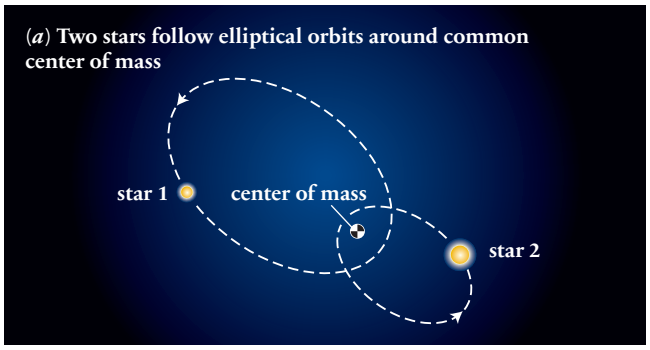
13.5 In part *a* of Example 13.1, what would the distance between the other person and you have to be for the gravitational force between the two of you to match the gravitational force exerted by Earth on you? Is it possible to verify your prediction?

13.2 Gravity and angular momentum

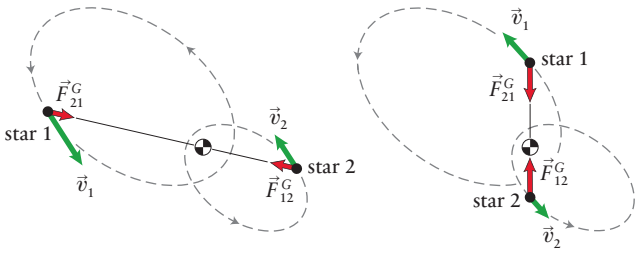
The force of gravity is a **central force**—a force whose line of action always lies along the straight line that connects the two interacting objects. A planet orbiting the Sun, for example, is subject to a force that is always directed toward the Sun. The fact that the force is central has an important consequence for angular momentum.

For example, consider an isolated system of two gravitationally interacting stars of masses m_1 and m_2 (**Figure 13.9**). We choose a coordinate system such that the origin is at the center of mass of the system. Because the system is isolated, the center of mass remains fixed at the origin. This means that the ratio r_1/r_2 of the objects' distances from the center of mass must remain the same regardless of the objects' motion (when $\vec{r}_{cm} = \vec{0}$, $m_1\vec{r}_1 + m_2\vec{r}_2 = \vec{0}$ and so $r_1/r_2 = m_2/m_1$; see Section 6.6). Note how the ratio r_1/r_2 in Figure 13.9 remains constant and how the orbits of the two stars have the same shape (scaled by the factor r_1/r_2).

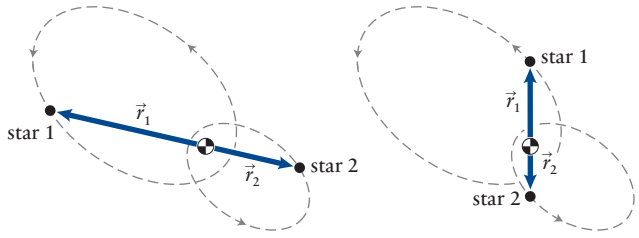
Figure 13.9 Gravitation is a central force, as illustrated by two stars orbiting their common center of mass.



(*b*) Gravitational forces stars exert on each other point along straight line connecting star centers. (Stars are shown at two instants in orbit.)



(*c*) Ratio r_1/r_2 of distances between each star and center of mass remains constant.



Because the center of mass always lies on the line of action of the gravitational forces, the torques caused by each of these forces about the center of mass are zero and therefore the angular momentum of each star about the center of mass doesn't change. So, not only is the angular momentum of the two-star system constant, but the angular momentum of each star about the center of mass is constant, too.

In an isolated system of two objects interacting through a central force, each object has a constant angular momentum about the center of mass.

Often we consider situations in which one of the two interacting objects is much more massive than the other, as shown in **Figure 13.10** on the next page. In that case, the center of mass lies nearly at the center of the more massive object. This permits us to view the situation from the position of the more massive object, which we can consider fixed at the origin. In Figure 13.10*a*, for example, the Earth moves in a nearly circular orbit about the Sun. Because the Sun is so much more massive than Earth, the center of the Sun is very close to the center of mass of the Sun-Earth system. We therefore can ignore the small difference and consider the two centers to coincide. If we consider the orbit to be circular, the center of mass also coincides with the center of the orbit. Figure 13.10*b* shows a comet moving along an elliptical orbit under the influence of a gravitational force exerted by the Sun. The center of mass again lies at the center of the Sun, which is at one of the foci of the ellipse (not at the center of the ellipse).

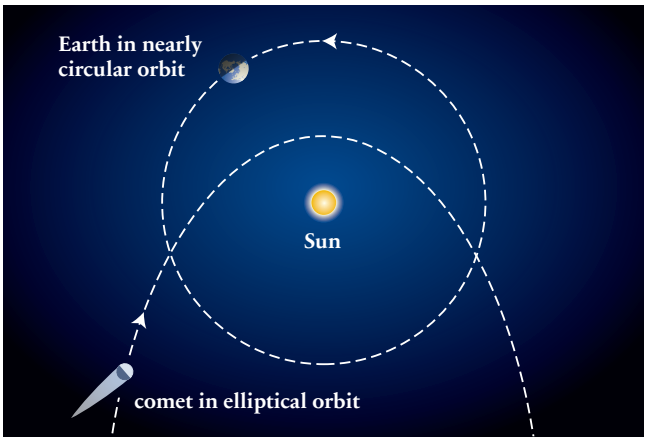
13.6 (*a*) If we ignore the influence of other celestial objects, the Earth-Sun system is isolated, and so it must rotate about the system's center of mass. Using the data in Table 13.1, determine how far from the center of the Sun the center of mass of the system is and what fraction of the Sun's radius this distance is. (*b*) As it passes the point shown in Figure 13.10*b*, does the comet speed up, slow down, or travel at constant speed? If the trajectory shown is part of an ellipse, where is the comet's velocity greatest? Where is it smallest?

Because the force is central and because the center of mass is at the Sun in Figure 13.10, the angular momentum of each of the orbiting objects about the Sun is constant. The meaning of this statement for a

CONCEPTS

Figure 13.10 Whether an orbit is circular or elliptical, the gravitational force on the orbiting body points toward the Sun. If the orbit is elliptical, the Sun is at one focus, not at the center of the ellipse.

(*a*) Orbits of Earth and hypothetical comet (not to scale)



(*b*) Force exerted by Sun on Earth and comet points toward Sun

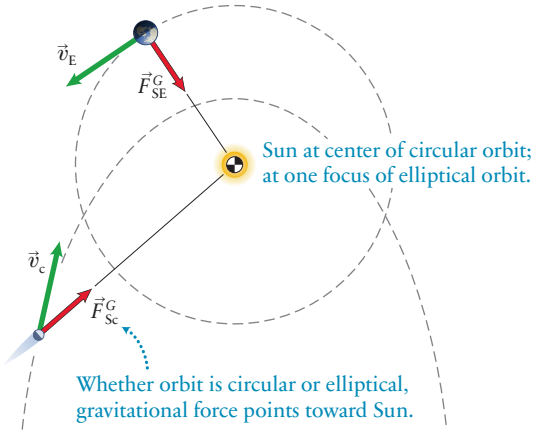
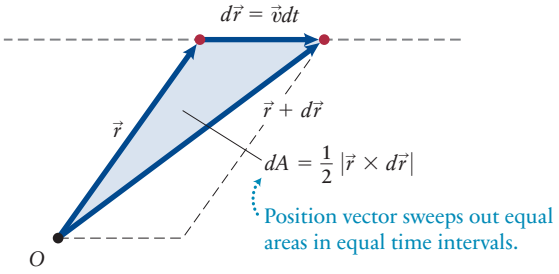
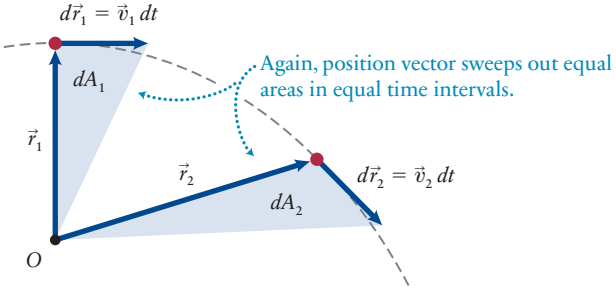


Figure 13.11 The position vectors of an object moving (*a*) with constant momentum along a straight line and (*b*) with constant angular momentum along an arbitrary path sweep out equal areas in equal time intervals.

(*a*) Object moving with constant momentum along straight line



(*b*) Object moving with constant angular momentum along curve



circular orbit is simple: If the angular momentum is constant, then the object's speed must be constant (that is, we are dealing with circular motion at constant speed).

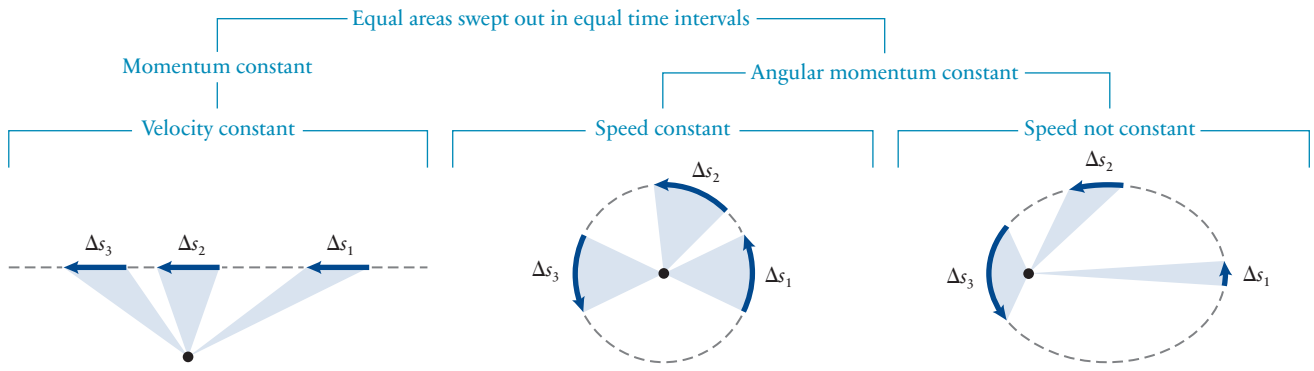
To understand the meaning for any other type of orbit, we must first introduce a geometric interpretation of angular momentum. Consider first the particle moving along a straight line illustrated in **Figure 13.11*a***. As we saw in Section 12.8 the angular momentum of any moving particle about the origin can be written as $\vec{L} = \vec{r} \times \vec{p}$, where $\vec{r}$ is the position of the particle (Eq. 12.43). In a time interval dt, the position vector $\vec{r}$ sweeps out the area of the shaded triangle in Figure 13.11*a*. As we saw in Section 12.8, the magnitude of the vector product of two vectors is equal to the area of the parallelogram defined by them. So $|\vec{r} \times d\vec{r}\,|$ is the area of the dashed parallelogram in Figure 13.11*a*—twice the area dA of the shaded triangle. Because the particle's displacement can be written as $d\vec{r} = \vec{v}\,dt$, we have $dA = \frac{1}{2}|\vec{r} \times \vec{v}|dt = \frac{1}{2}L/m\,dt$. The rate at which the area A is swept out by the position vector is thus $dA/dt = \frac{1}{2}L/m$. In words:

> **The angular momentum of a particle about an origin is proportional to the rate at which area is swept out by the particle's position vector.**

So, when the angular momentum of a particle is constant, the rate dA/dt is also constant. Figure 13.11*b* shows the motion of an object along a curved trajectory. In two equal time intervals dt, the position vector sweeps out areas dA_1 and dA_2. If the angular momentum of the object about O is constant, then $dA_1 = dA_2$. The motion caused by a central force directed to the origin is such that the radius vector from the origin sweeps out equal areas in equal time intervals *regardless of the shape of the trajectory*. In the early 17th century, before Newton formulated the law of gravity and long before the concept of angular momentum was introduced, Johannes Kepler had already deduced from astronomical observations that the straight line from the Sun to each planet sweeps out equal areas in equal time intervals. (See the box "Kepler's laws of planetary motion.")

CONCEPTS

Figure 13.12 Three forms of motion with constant angular momentum.



(*a*) Motion at constant velocity along straight line

(*b*) Motion at constant speed along circular path

(*c*) Motion about one focus of an ellipse in response to central force

Figure 13.12 shows three types of motion with constant angular momentum: motion at constant speed along a straight line (no force; constant velocity), circular motion at constant speed (central force; constant speed), and elliptical motion (central force; constant rate of area swept). For the third it is not immediately obvious that the angular momentum is constant: Neither the velocity direction nor the speed of the moving object is constant. For all three types of motion, however, the position vector sweeps out the same area in equal time intervals. To do so an object in elliptical motion (Figure 13.12*c*) must move fastest when it is closest to the object it is orbiting and slowest when it is the farthest away, in agreement with the answer to Checkpoint 13.6, part *b*.

EXAMPLE 13.2 Pluto

Of all the objects in Table 13.1, Pluto has the orbit with the greatest eccentricity. (*a*) What is the ratio of the orbit's semiminor axis b to its semimajor axis a? (*b*) What is the ratio of Pluto's speed at perihelion to its speed at aphelion?

❶ **GETTING STARTED** I begin by making a sketch of Pluto's orbit, using the Sun as one focus (**Figure 13.13** on the next page). From Figure 13.8 I know that the orbit is very nearly circular but the Sun is off-center, making the orbit slightly elliptical. I mark the perihelion (shortest Sun-orbit distance) and aphelion (longest Sun-orbit distance) and show velocity vectors at those locations because I am asked about speeds there. In the ellipse as I have oriented it, with the line defining the foci horizontal, the semiminor axis b extends from the

(Continued)

Kepler's laws of planetary motion

The German mathematician and astronomer Johannes Kepler devoted much of his life to analyzing celestial motion data collected by Tycho Brahe. Kepler discovered three laws of planetary motion.

Kepler's first law describes the shape of the planetary orbits:

> **All planets move in elliptical orbits with the Sun at one focus.**

Although the deviation from circular orbits is small, this statement was a radical departure from the accepted wisdom, dating back to Plato, that the planets, being heavenly bodies, were perfect and therefore could move in only perfect circles or combinations of circles.

Kepler's second law reveals that, even if planets are not in circular motion at constant speed, their motions obey the following requirement:

> **The straight line from any planet to the Sun sweeps out equal areas in equal time intervals.**

Kepler's third law relates the planetary orbits to one another:

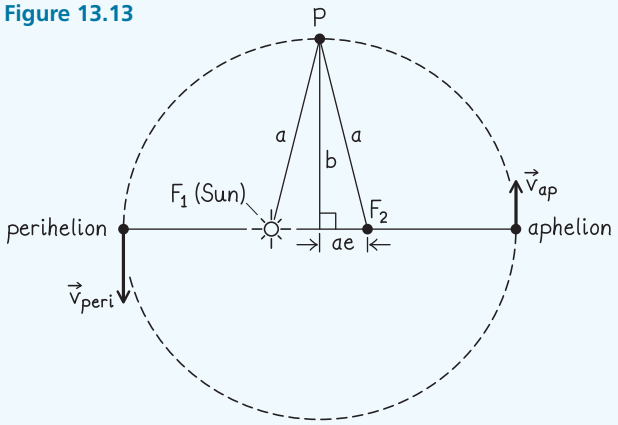
> **The squares of the periods of the planets are proportional to the cubes of the semimajor axes of their orbits around the Sun.**

Kepler discovered this third law by painstakingly examining, over a period of many years, countless combinations of planetary data.

In keeping with Aristotelian notions, Kepler believed that a force was necessary to drive the planets along their orbits, not to keep them in orbit. Consequently, Kepler was unable to provide a correct explanation for these three laws. It was not until Newton that the single unifying reason for these laws was established. As we've seen, Kepler's three laws follow directly from the law of gravity: The first and third laws are a consequence of the $1/r^2$ dependence, and the second law reflects the central nature of the gravitational force.

CONCEPTS

Figure 13.13



center of the major axis to point P at the top of the ellipse. The ratio of the semimajor and semiminor axes can be obtained from either right triangle in Figure 13.13. Because the sum of the distances from any point on an ellipse to the two foci is $2a$ (see Figure 13.7), I know that $\overline{F_1P} = \overline{F_2P} = a$.

2 DEVISE PLAN For part *a*, I can use the Pythagorean theorem and the values in Table 13.1 to calculate the ratio b/a. For part *b* I know that the planet's angular momentum about the Sun must be constant because the force exerted by the Sun on Pluto is a central force. I can therefore calculate the angular momentum at perihelion and at aphelion and set the two equal to obtain the ratio $v_{\text{peri}}/v_{\text{ap}}$.

3 EXECUTE PLAN (*a*) Applying the Pythagorean theorem to the right triangle containing F_2 gives $b = +\sqrt{a^2 - (ae)^2} = a\sqrt{1 - e^2}$. Using the value of e from Table 13.1, I get $b/a = +\sqrt{1 - (0.25)^2} = 0.97$. ✔

(*b*) At perihelion, the Sun-Pluto distance is $(a - ae)$, and the angular momentum is $L_{\text{peri}} = m(a - ae)v_{\text{peri}} = ma(1 - e)v_{\text{peri}}$ (Eq. 11.36); at aphelion, $L_{\text{ap}} = ma(1 + e)v_{\text{ap}}$. Because $L_{\text{peri}} = L_{\text{ap}}$, I get

$$ma(1 - e)v_{\text{peri}} = ma(1 + e)v_{\text{ap}},$$

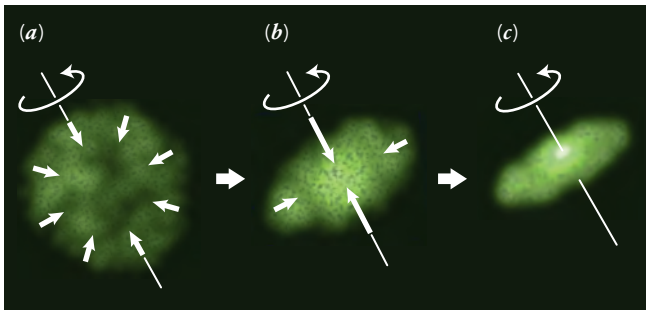
so

$$\frac{v_{\text{peri}}}{v_{\text{ap}}} = \frac{1 + e}{1 - e} = \frac{1 + 0.25}{1 - 0.25} = 1.7. \checkmark$$

4 EVALUATE RESULT The semiminor and semimajor axes are equal to within 3%, which agrees with what is shown in Figure 13.8. That v_{peri} is greater than v_{ap} agrees with my answer to Checkpoint 13.6, part *b*.

Conservation of angular momentum is also responsible for the flat, disk shape of most galaxies and of our solar system. **Figure 13.14** illustrates the gravitational contraction of a slowly rotating cloud of interstellar gas and dust. Because the cloud's rotational inertia decreases as it contracts, its rate of rotation must increase, just as a skater spins faster with her arms pulled in. In the plane perpendicular to the rotation axis, part of the gravitational force directed toward the center of the cloud goes into providing the centripetal acceleration. Regions near this plane therefore undergo less contraction than do regions along the axis. This differential contraction results in the characteristic disk shape of many galaxies.

Figure 13.14 (*a*) Gravitational contraction of a slowly rotating cloud of interstellar gas and dust. (*b*) Conservation of angular momentum requires the rotational velocity to increase as the cloud contracts and its rotational inertia decreases. (*c*) Contraction is less efficient in the plane perpendicular to the rotation axis because part of the gravitational interaction goes into supplying the centripetal acceleration.



13.3 Weight

The law of gravity states that the force of gravity depends on location. **Figure 13.15** shows astronauts orbiting Earth at an altitude of about 400 km. Are they floating "weightlessly" inside their spacecraft because the force of gravity at their location has been reduced to a negligible value? The calculation is easy and worth doing, as the next checkpoint shows.

13.7 The space shuttle typically orbits Earth at an altitude of about 300 km. (*a*) By what factor is the shuttle's distance to the center of Earth increased over that of an object on the ground? (*b*) The gravitational force exerted by Earth on an object in the orbiting shuttle is how much smaller than the gravitational force exerted by Earth on the same object when it is sitting on the ground? (*c*) What is the acceleration due to gravity at an altitude of about 300 km? (*d*) While in orbit, the shuttle's engines are off. Why doesn't the shuttle fall to Earth?

After answering this checkpoint, you may be wondering in what sense objects inside the space shuttle are "weightless." To resolve this issue, we must define what we mean by *weight*.

Figure 13.15 Astronauts Shane Kimbrough and Sandra Magnus float "weightlessly" in Space Shuttle Endeavour during mission STS-126.



CONCEPTS

Figure 13.16 Two methods for weighing a brick.

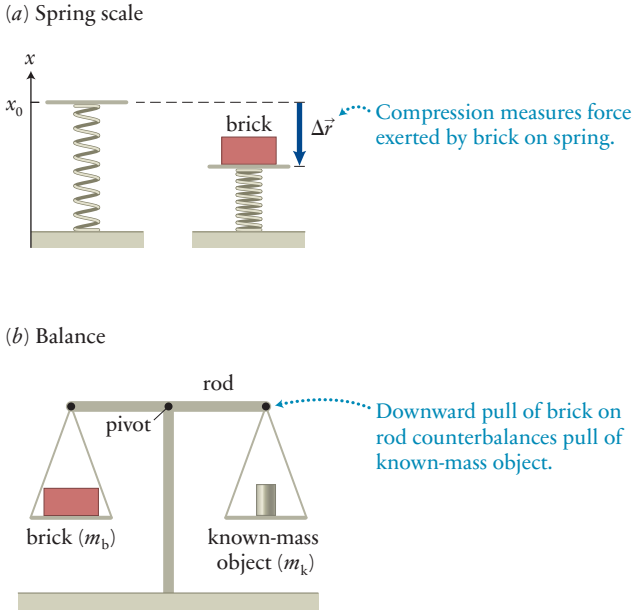


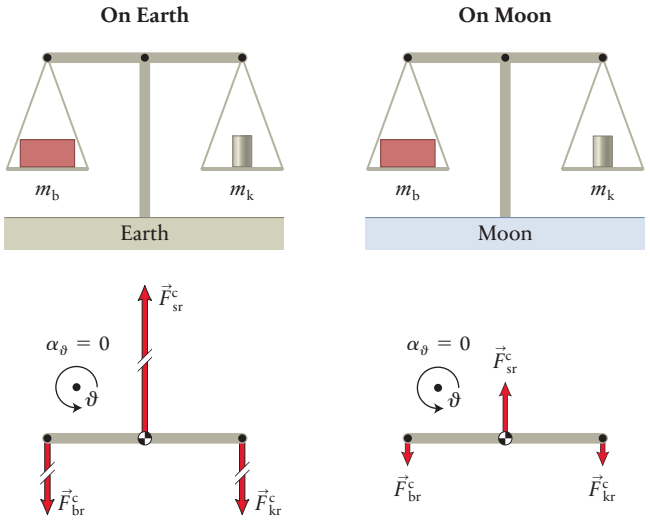
Figure 13.16 illustrates two types of devices used for weighing objects. When a load is put on a **spring scale,** the load either compresses or stretches a spring inside the scale (Figure 13.16*a*). The change in the spring's length is proportional to the force exerted on it by the load: $F^c_{ls\,x} = k(x - x_0)$ (Eq. 8.18). The deformation of the spring therefore provides a measure of the downward force exerted on it by the load. When you stand on a bathroom scale, for example, the downward force exerted by your body on the scale compresses a spring inside it, which, in turn, rotates a dial, and you read your weight off the dial. The rotation of the dial is proportional to the force exerted by your body on the scale. (The scale was calibrated at the factory by putting objects of known mass on it.)

13.8 A 1.0-kg brick is placed on a spring scale. (*a*) What is the magnitude of the force exerted by Earth on the brick? (*b*) What is the magnitude of the force exerted by the brick on the scale? (*c*) How many steps are required to answer part *b*? (*d*) Does the spring scale measure the pull of gravity on the brick?

Another type of weighing device, the **balance,** illustrated in Figure 13.16*b*, works as follows. A brick whose mass m_b is to be determined is suspended from one end of a rod pivoted about its center, and objects of known masses are suspended from the other end until the rod is in mechanical equilibrium. In mechanical equilibrium, the downward forces exerted by the loads on the two ends of the rod are equal, and so m_b must be equal to the sum m_k of the known masses.

Don't both devices measure the same thing? You may be surprised to learn that the answer to this question is *no*. To see why, suppose we repeat the experiments

Figure 13.17 Using a balance at Earth's surface and on the Moon. If the rod is balanced on Earth, the same objects balance the rod on the Moon.



illustrated in Figure 13.16 on the surface of the Moon. The gravitational force exerted by the Moon on an object sitting on the lunar surface is about six times smaller than the gravitational force exerted by Earth on the same object sitting on Earth's surface. If you were carrying a heavy load on the Moon, you would discover that the force needed to support it is less than on Earth; on the Moon the load feels a lot lighter. As the extended free-body diagrams in **Figure 13.17** show, however, the smaller gravitational pull of the Moon does not affect the mechanical equilibrium of the balance: The downward pull diminishes on *both* sides, and so the rod remains in mechanical equilibrium. With the balance, we still obtain $m_b = m_k$, just as we do on Earth.

The reading of the spring scale on the Moon, however, is not the same as it is on Earth. Because of the weaker pull of gravity on the Moon, the deformation of the spring is smaller, as shown in **Figure 13.18** on the next page. Spring scale readings on the Moon are about six times smaller than on Earth. So the spring scale tells us that the load weighs less on the Moon than on Earth, while the balance tells us that it weighs the same. This difference arises because:

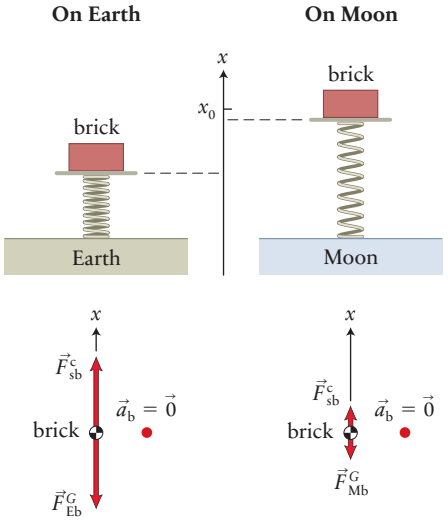
A spring scale measures the downward force exerted on it by its load, but a balance compares gravitational forces and measures mass.

The result obtained with a spring scale depends on the strength of the gravitational pull on the load; the result obtained with a balance does not depend on the strength of the gravitational pull.*

*Many modern weighing devices are called "balances" even though in reality they operate like spring scales. Typically, these devices employ a pivoted rod, but instead of comparing gravitational forces, one end of the rod is connected to a spring or other mechanism for generating a calibrated force.

CONCEPTS

Figure 13.18 Free-body diagrams for a brick on a spring scale on the surface of Earth and the Moon. Because of the smaller gravitational attraction by the Moon, the spring's compression on the Moon is less than it is on Earth.

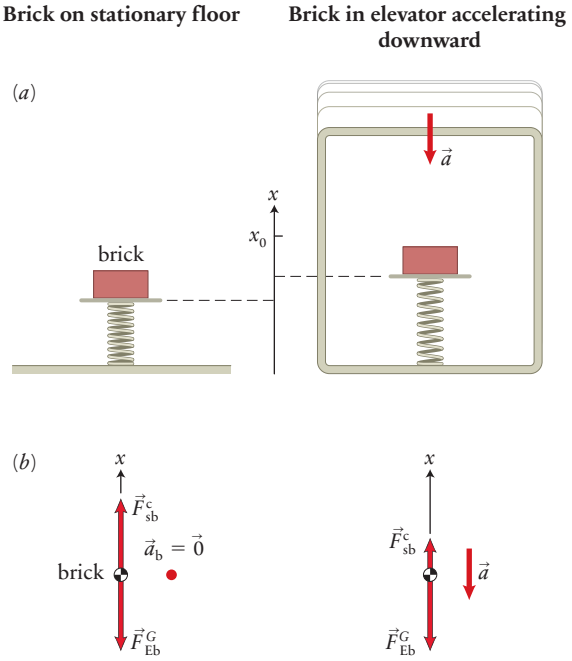


13.9 A 1.0-kg brick is placed on a spring scale inside the space shuttle orbiting Earth at 300 km altitude. (*a*) What is the magnitude of the force exerted by Earth on the brick? (*b*) What is the magnitude of the force exerted by the brick on the spring scale? (Hint: See the astronauts in Figure 13.15.) (*c*) Does the spring scale measure the pull of gravity on the brick? (*d*) Why is your answer different from your answer to Checkpoint 13.8, part *d*?

The answers to this checkpoint show that the reading of a spring scale depends not only on the gravitational pull on its load but also on the acceleration of the scale. To see this point more clearly, consider the situation shown in **Figure 13.19***a*. Two bricks of identical mass m are placed on two identical spring scales, one on solid ground and the other in an elevator that is accelerating downward. Figure 13.19*b* shows the free-body diagrams for both bricks. Because the brick in the elevator has a downward acceleration, we know that the vector sum of the forces exerted on it has to be nonzero and directed downward. The downward force of gravity has not changed, however. Consequently, the upward force exerted by the scale on the brick must be *smaller* than the magnitude of the downward force of gravity acting on it. Because the force exerted by the scale on the brick and the force exerted by the brick on the scale form an interaction pair, we conclude that the force exerted by the brick on the scale in the elevator is smaller than that exerted by the identical brick on the stationary spring scale. The spring scale in the elevator therefore gives a lower reading than the stationary scale.

If you are surprised that the result of weighing an object with a spring scale depends on acceleration, try the following experiment. Put an object in the palm of your hand with your arm stretched out in front of you,

Figure 13.19 A spring scale holding a brick gives a different reading when it is stationary and when it is in an elevator accelerating downward. Because of the acceleration of the elevator, $\Delta x_{elev} < \Delta x_{stat}$.



then rapidly move your hand downward. While you accelerate your hand downward, you can feel that the object presses less on your hand than it does when you hold your hand still. In other words, while your hand accelerates downward, the object in it feels lighter—it feels as though it "weighs" less.*

13.10 (*a*) An object is placed on a spring scale in an elevator moving upward at a constant speed. Is the reading of the scale greater than, smaller than, or the same as the reading obtained in a stationary elevator? (*b*) How does the reading of a scale in an elevator that has upward acceleration $a = 0.5g$ compare with the reading obtained in a stationary elevator? (*c*) What if the elevator has downward acceleration $a = g$?

As this checkpoint shows, objects in an elevator that is in free fall produce a zero spring scale reading. No force is necessary to support them: They, too, fall freely along with the elevator. Inside a freely falling elevator, a spring scale under your feet—falling away underneath you, together with the elevator and everything inside it—reads zero. If you were to hold a briefcase in the freely falling elevator, you would discover that the briefcase does not need to be supported by your hand because the force of gravity gives it the same

*You can also try this by accelerating your hand upward, in which case the object feels heavier. The effect, however, is less pronounced because it is more difficult to accelerate your arm upward. In the downward direction, gravity and your muscles work together to provide the acceleration; in the upward direction, your muscles have to not only accelerate your arm but also counteract the effect of gravity.

CONCEPTS

downward acceleration g as you. The briefcase feels weightless—you could even let go and the briefcase would ue to hover next to you.

Any object in free fall—that is, any object subject to only a force of gravity—experiences weightlessness.

If this appears odd, consider once again the astronauts in orbit in Figure 13.15. The only force exerted on them is the force of gravity, and so they *are* in free fall! This free fall is precisely what makes them weightless. Because of their large speed and the curvature of Earth, however, they do not come closer to the ground. As they fall, the ground below them recedes at the same speed as they fall (**Figure 13.20**).

13.11 (*a*) How far does the space shuttle in Figure 13.20, 300 km above Earth, fall in 1.0 s? (*b*) If the radius of Earth is 6400 km, what is the shuttle's speed?

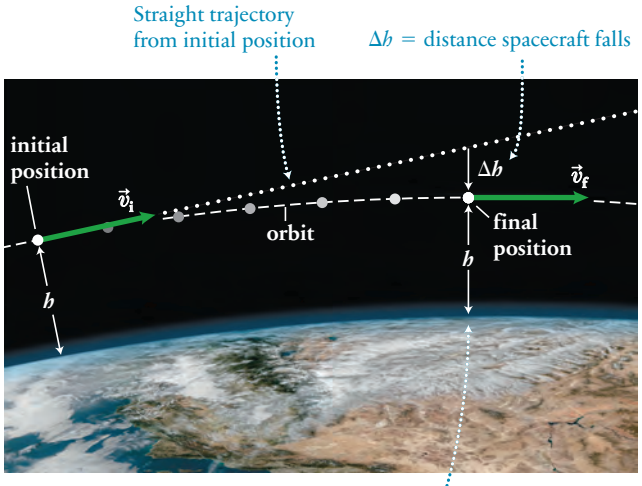
You don't need to be an astronaut to experience weightlessness. When you jump off a chair, you are momentarily weightless; during your short free fall, an object in the palm of your hand hovers weightlessly above your hand. A plane flying across a large parabolic arc may be in a controlled free fall for short time intervals. As shown in **Figure 13.21**, people inside such an airplane float weightlessly, just as astronauts in orbit do.

13.12 Planes can be put (safely) into free fall for periods up to 40 s. (*a*) If the plane shown in Figure 13.21 was flying horizontally before being put into free fall, how much altitude does it lose in those 40 s? (*b*) What is the plane's final vertical speed of descent? (*c*) At the instant shown, is the vertical component of the plane's velocity directed upward, directed downward, or zero?

The discussion in this section shows that our notion of *weight* corresponds most closely to the reading of a spring scale: The perceived "heaviness" of an object depends, as does the reading on the scale, not only on the pull of gravity on the object but also on its acceleration. The "weightless" astronauts in the space shuttle would indeed register a zero scale reading, even though the pull of gravity on them is almost identical to the pull at Earth's surface.

Because spring scale readings depend on acceleration, however, weight is often defined as the force of gravity exerted *on* an object. Spring scale readings are then called *apparent* or *effective weight*. With this definition, the astronauts in Figure 13.15 and the people in Figure 13.21 are not weightless because the force of gravity on them (their "weight") is the same as it is on Earth's surface, and so they must be called *apparently weightless*. To avoid all these difficulties and the conflict between our intuitive notion of weight (spring scale reading) and the common definition of weight (force of gravity), we therefore avoid using the word *weight* in this book. Instead we shall use the more precise "force of gravity" to refer to the downward pull of gravity on an object and "scale reading" to refer to the force exerted by an object on a spring scale.

Figure 13.20 As a spacecraft orbits Earth a distance h above the ground, it falls a distance Δh. Because the craft moves so fast in a direction tangential to the (curved) ground, the ground "falls away" at the same rate, and the distance between the craft and the ground remains constant.



Height h of orbit remains constant because of distance Δh spacecraft falls.

13.13 Consider an airplane flying at a constant speed of 900 km/h, 10 km above the ground. (*a*) What is the plane's downward acceleration? (*b*) Draw a free-body diagram for the airplane. (*c*) Why don't people in the airplane feel weightless as astronauts do traveling 290 km higher up?

Figure 13.21 If an airplane briefly follows a parabolic path such that it has a downward acceleration g, the airplane and the people inside it are in free fall. In this picture, astronaut candidates aboard a KC-135 Stratotanker briefly experience weightlessness as they and the plane are in a short free fall.



CONCEPTS

13.4 Principle of equivalence

Let us return to the remarkable equality of mass and inertia. The simplest possible test of this equality is obtained by dropping two objects that have different inertias and are made of different materials and measuring their accelerations. As Galileo discovered, *all* objects fall with the same acceleration g. It is true that Earth pulls twice as hard on an object that has twice the mass, but because this object also has twice the inertia, it doesn't accelerate any faster. The equality of mass and inertia has been tested experimentally to one part in 10^{12}.

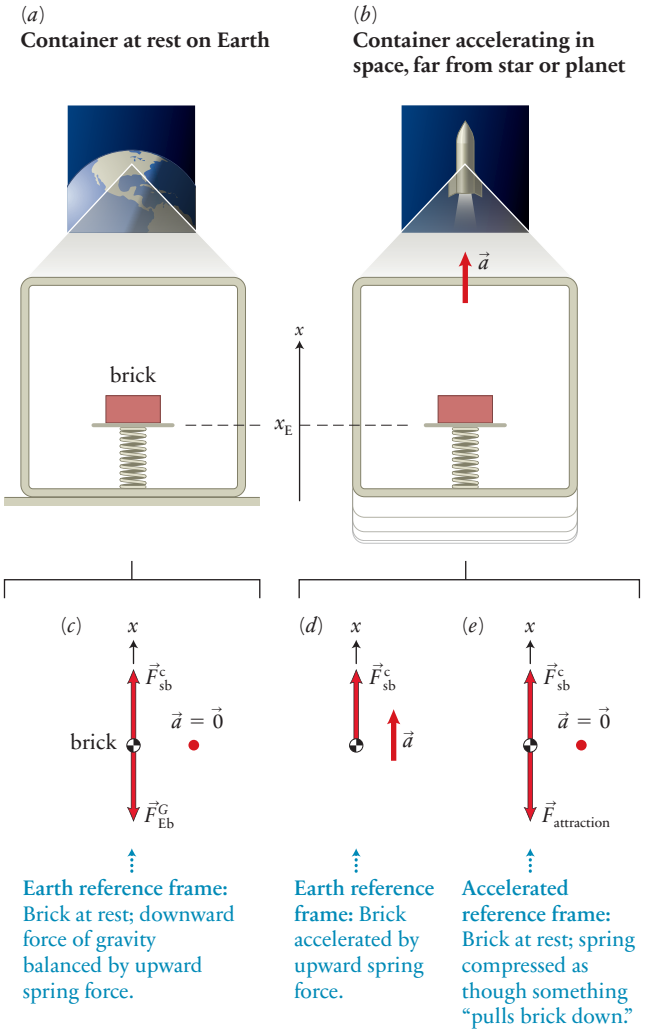
The equality of mass and inertia may seem obvious because we are so accustomed to the fact that heavier objects not only take more effort to support against gravity but also take more effort to be set in motion. Intuitively, both properties simply go by the amount of matter in the object. Yet this equality suggests a much deeper connection between gravity and acceleration.

In Section 13.3 we encountered another manifestation of this equivalence between gravity and inertia: The reading on a spring scale depends not only on the gravitational pull on the load but also on the acceleration of the scale. We can change the reading of the scale—our notion of weight—simply by accelerating the scale. This raises an interesting question: Can we design an experiment that distinguishes between gravity and an accelerated reference frame?

Consider a brick placed on a spring scale inside a container at rest on the surface of Earth, as illustrated in **Figure 13.22*a***. Because of the gravitational attraction exerted by Earth on the brick, the brick compresses the spring to a scale reading x_E. To an observer in the Earth reference frame the brick is at rest, and so the upward force exerted by the scale is equal in magnitude to the downward force exerted by Earth (see Figure 13.22*c*). Now picture the same arrangement in space, far away from any planet or star, where the influence of gravity is negligible and so we can say that $g = 0$. Suppose the container is towed by a spaceship; if the spaceship fires its rockets, the container accelerates, as illustrated in Figure 13.22*b*. Objects inside the container no longer float weightlessly, but being dragged along by the container, they rest against the bottom of the container. If the brick is placed on the scale, the scale registers a reading proportional to the force exerted by the spring on the brick. This force is required to give the brick the same acceleration $\vec{a}$ as the container. From the point of view of an observer in the Earth reference frame, the brick has an upward acceleration $\vec{a}$ (see Figure 13.22*d*).

If the rockets are fired such that $a = g$, could you, from inside the container, tell the difference between being accelerated in space and being on Earth? To you, the scale and the scale reading appear just as they do on Earth. An object released from your hand falls with an acceleration g, just as it does on Earth. From inside the container, with both of your feet solidly on the floor, everything looks and feels just as it does on Earth—all objects appear to be attracted to the container floor.

Figure 13.22 If you observe a brick compressing a spring scale in an enclosed container, you cannot tell whether the compression is due to a gravitational interaction or to acceleration of the container.



Having observed all this, you draw the free-body diagram shown in Figure 13.22*e*: From your point of view, the brick is at rest, and so the vector sum of the forces exerted on it must be zero. So, in addition to an upward force exerted by the compressed spring on the brick, you draw a downward force, equal in magnitude to that of the upward force. To you, it appears as if this force is due to a downward attraction experienced by all objects inside the container. Without peeking outside, you cannot determine whether this attraction is present because underneath the container there is a planet or because the container is accelerating upward in space.

Notice that the fictitious attractive force experienced by objects inside the container is no less mysterious than the force of gravity. Both forces have the same effect: Objects are pulled toward the bottom of the container according to their mass or inertia (the distinction has now disappeared). Although the gravitational force

CONCEPTS

A flight simulator for the Sukhoi Superjet 100.



is more familiar, and although Newton's law of gravity describes its behavior, we really have no idea how gravity "works." The conclusion is thus:

> **One cannot distinguish locally between the effects of a constant gravitational acceleration of magnitude g and the effects of an acceleration of the reference frame of magnitude $a = g$.**

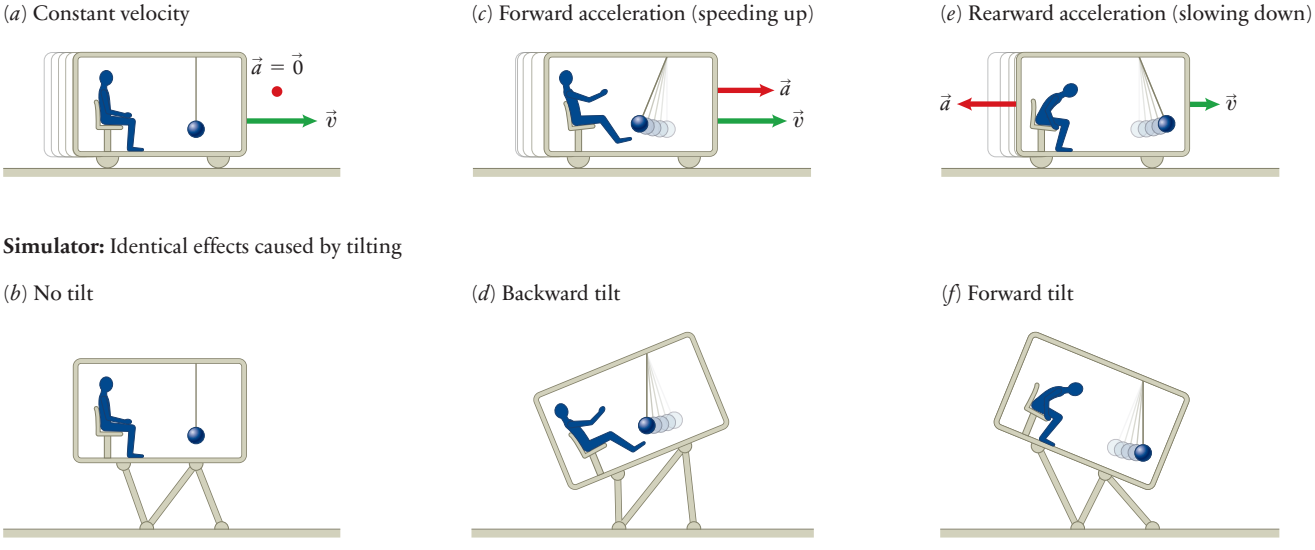
This statement, called the **principle of equivalence,** is the starting point for the theory of general relativity developed by Albert Einstein in the early 1900s. The word *locally* is necessary because on a large scale, the $1/r^2$ dependence of the gravitational force becomes apparent: For a very tall container in Figure 13.22, the gravitational acceleration at the top of the container would be smaller than at the bottom; for a very wide container, it would become noticeable that objects are falling not parallel to one another but toward a single point—the attracting object's center. In the theory of general relativity, the effects of gravity are explained in an entirely geometric way. Gravity and acceleration are one.

On a practical note, our inability to distinguish between gravity and acceleration is exploited in aircraft flight simulators and motion simulators in the amusement industry. These devices are programmed to tilt in concert with a film simulating a ride in a vehicle, airplane, or spaceship. The effect is a startlingly realistic sensation of motion. The principle of operation is shown in Figure 13.23. Suppose you are in a closed container, with a ball suspended from the ceiling. If the container rolls at constant speed along a track, as in Figure 13.23*a*, you cannot *feel* the motion and the string hangs vertically, although sounds from the wheels and little bumps in the track might give away that the container is in motion. The law of inertia (see Section 6.2) tells us, however, that you cannot determine the difference between moving at constant speed and sitting at rest. So, with the container on a stationary base, as in Figure 13.23*b*, if the container shakes a bit to simulate bumps and you hear the same sounds and see a movie of a landscape that flows by at constant speed, the illusion of moving at constant speed is complete.

Now suppose the container speeds up, as in Figure 13.23*c*. The chair presses against your back to accelerate you along with the container, and as a result you feel pushed backward against the chair. As Figure 13.23*d* shows, the same effect can be obtained by tilting a stationary container backward (together with appropriate sounds and visual clues). In both cases, also, the ball suspended from the ceiling tilts toward you, and objects don't fall straight to the floor. From inside the container, you cannot tell whether these things happen because the container accelerates or because gravity is in a different direction.

CONCEPTS

Figure 13.23 The principle of equivalence states that an observer inside a closed container cannot distinguish between motion at constant speed and rest, between speeding up and tilting backward, or between slowing down and tilting forward.



To simulate slowing down (see Figure 13.23*e*), the container needs to be tilted forward (see Figure 13.23*f*). The result is that you feel pushed forward, as in a sharply braking car, and the ball tilts away from you. Again you cannot tell whether the container is slowing down or gravity is pulling in a different direction.

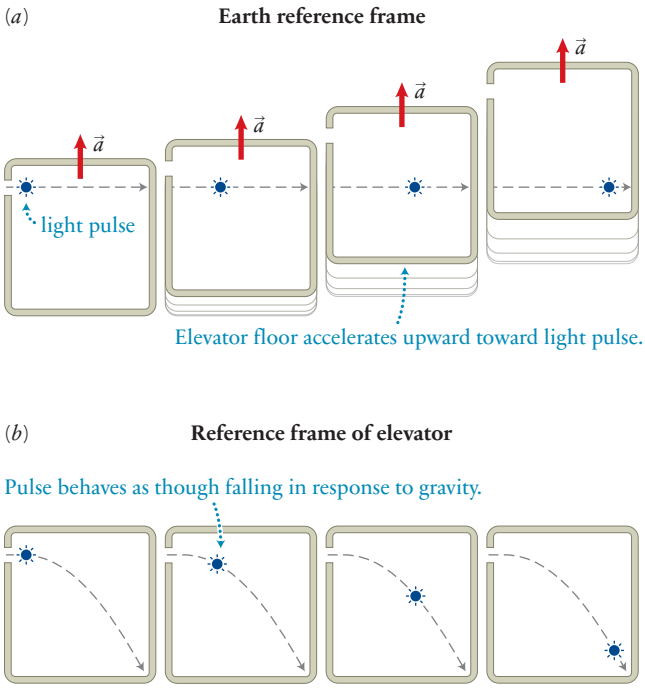
13.14 How should the simulator in Figure 13.23 be tilted to simulate a right turn?

The simulator example may give you the impression that the principle of equivalence has more to do with our senses than with the properties of the universe. This is not so, however. Because anything that moves along a straight line in an inertial reference frame must move along a curved path in an accelerated reference frame, the principle of equivalence tells us that *anything* that moves near an object that has mass must move along a curved path. **Figure 13.24** shows that a light pulse that travels along a straight line in the Earth reference frame but then enters an accelerating elevator must move along a curved path as viewed from inside the elevator. This effect is purely kinematic, and so given the equivalence between gravity and acceleration, the bending should also occur if the elevator, instead of accelerating upward, were placed on top of an object with a large mass.

13.15 Light travels at approximately 3.0×10^8 m/s. (*a*) How long does it take for a light pulse to cross an elevator 2.0 m wide? (*b*) How great an acceleration is necessary to make the pulse deviate from a straight-line path by 1.0 mm? Is it likely that this effect can be observed? (*c*) If light is bent by the gravitational pull of an object, light should "fall" when traveling parallel to the surface of Earth. How far does a beam of light travel in 0.0010 s, and how much does it fall over that distance? Is it likely that this effect can be observed?

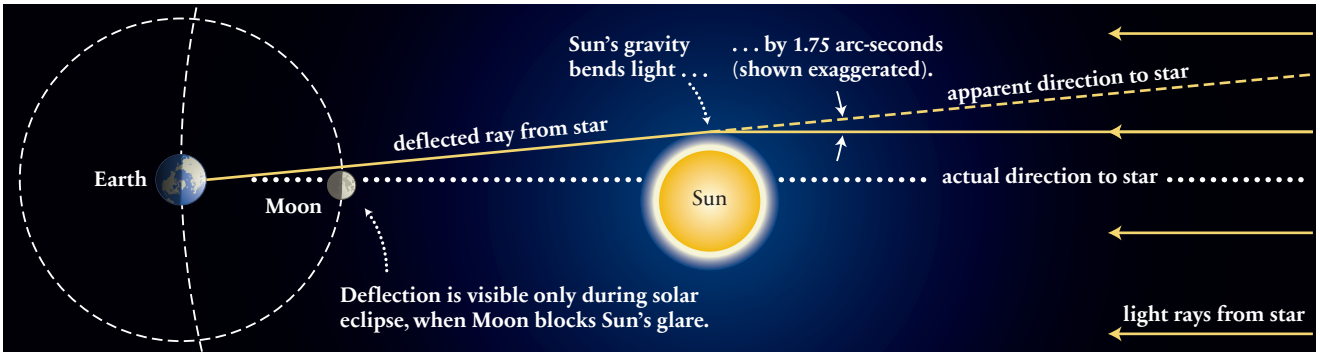
Although the calculations in this checkpoint show that detection of the gravitational bending of light is difficult, the effect was observed within a few years of its prediction in 1915. Because of the large mass of the Sun, a ray that passes close to the Sun should undergo a deflection. Although it is normally impossible for us earthlings to see any stars close to the Sun, the light from the Sun is blocked by the Moon during a solar eclipse (**Figure 13.25**). By comparing a picture of stars near the Sun taken during an eclipse with a picture of the same stars taken when the Sun is far away from them (in other words, at night), it is possible to observe the bending of light caused by the Sun's gravitational attraction. The measured values of such deflections are in excellent agreement with the theoretical values.

Figure 13.24 A light pulse traverses an accelerating container.



13.16 (*a*) What would the scale reading be if the container in Figure 13.22*b* were traveling at constant speed instead of accelerating? (*b*) Describe the path of the light pulse in the reference frame of the elevator in Figure 13.24*b* if the elevator were traveling at constant speed.

Figure 13.25 When the Sun passes close to our line of sight to a star, the Sun's gravity bends the star's light, slightly shifting the star's apparent location in the sky. This effect is visible only during a solar eclipse, when the Moon blocks the Sun's glare. Diagram not to scale.



CONCEPTS

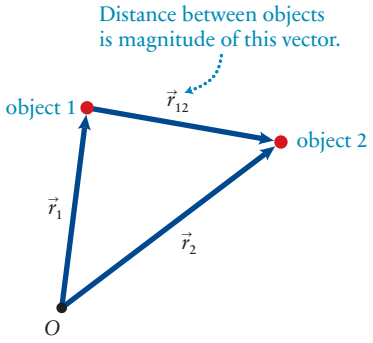
Self-quiz

1. Determine the ratio of Earth's gravitational force exerted on a 1.0-kg object at the following positions to that at sea level: (*i*) the top of Mt. Everest, altitude 8848 m; (*ii*) a shuttle orbiting at 200 km; (*iii*) the orbit of the Moon. Is the 1.0-kg object weightless in any of these circumstances?
2. The orbital period of Jupiter around the Sun is 11.86 times that of Earth. What is the ratio of their distances to the Sun?
3. If the speed of a planet in its orbit around the Sun changes, how can the planet's angular momentum be constant?
4. A brick of unknown mass is placed on a spring scale. When in equilibrium, the scale reads 13.2 N. Which of the following statements is/are true? (*i*) Earth always exerts a gravitational force of 13.2 N on the brick. (*ii*) The normal force the scale exerts on the brick is 13.2 N. (*iii*) The brick has an inertia of 1.32 kg.
5. Draw a free-body diagram of an individual experiencing weightlessness in an elevator. What must the elevator's acceleration be for this to occur?
6. If you were to travel in a vertical circle, at what point in the circle would you be most likely to experience weightlessness?

Answers

1. The gravitational force is proportional to the inverse square of the distance between Earth and the object. For each case, we have $F_{\text{pos}}/F_{\text{sea level}} = r_{\text{sea level}}^2/r_{\text{pos}}^2$, where $r_{\text{sea level}} = 6.378 \times 10^6$ m (Earth's radius) and r_{pos} is the distance between the position in question and Earth's center. (*i*) The distance to Earth's center is the sum of Earth's radius and the altitude of Mt. Everest: $r_{\text{pos}} = (6.378 \times 10^6\ \text{m}) + (8848\ \text{m}) = 6.387 \times 10^6$ m; $F_{\text{pos}}/F_{\text{sea level}} = 0.997$. (*ii*) $r_{\text{pos}} = (6.378 \times 10^6\ \text{m}) + (200{,}000\ \text{m}) = 6.578 \times 10^6$ m; $F_{\text{pos}}/F_{\text{sea level}} = 0.940$. (*iii*) The distance to Earth's center is the radius of the Moon's orbit. Given that the Moon's orbit is nearly circular, I use the value of the semimajor axis in Table 13.1: $r_{\text{pos}} = 3.84 \times 10^8$ m; $F_{\text{pos}}/F_{\text{sea level}} = 0.000276$. The object is never weightless because, in each case, it experiences a gravitational force and, in all but the last case, the magnitude of this force is very close to the magnitude at sea level on Earth.
2. The period squared is proportional to the cube of the orbital radius. Therefore, the square of the ratio of the periods ($11.86^2 = 141$) is equal to the cube of the ratio of the orbital radii. The ratio of the radii is then $141^{1/3} = 5.2$. Check this value against the data in Table 13.1.
3. If the angular momentum is constant but the speed changes, the rotational inertia must change. When a planet's speed increases, it must move nearer the object it orbits (or its inertia must decrease) to keep the angular momentum constant. Conversely, when a planet's speed decreases, it must move away from the object it orbits (or its inertia must increase).
4. Only statement (*ii*) is true. The scale reading may change from location to location, and the inertia of the brick is 1.32 kg only when the gravitational acceleration at the specific location is 10 m/s^2.
5. If a person experiences weightlessness, the only force acting on her or him is the gravitational force. Contact with the elevator floor is not necessary. If the person is in free fall, the elevator must also be in free fall. The free-body diagram consists of just one force vector, $\vec{F}_{\text{Ep}}^{G}$ downward.
6. To experience weightlessness, the only force acting on you is the gravitational force exerted by Earth, which points downward. This downward force provides the centripetal acceleration needed for traveling in a circle. The most likely position from which the center of the circle is downward is the top of the circle.

Figure 13.26 The distance between two objects whose positions are given by the vectors $\vec{r}_1$ and $\vec{r}_2$ is equal to the magnitude of the vector $\vec{r}_2 - \vec{r}_1$.



13.5 Gravitational constant

As we saw in Section 13.1, Newton's **law of gravity** states that the gravitational force exerted by a object of mass m_1 on an object of mass m_2 is an attractive, central force that is proportional to the product of the masses and inversely proportional to the square of the distance between them. If the positions of the two objects are given by the vectors $\vec{r}_1$ and $\vec{r}_2$, then the distance between them is $r_{12} = |\vec{r}_2 - \vec{r}_1|$, as shown in **Figure 13.26**. We can thus write the magnitude of the gravitational force as

$$F_{12}^G = G\frac{m_1 m_2}{r_{12}^2}, \tag{13.1}$$

where the proportionality constant G is called the **gravitational constant.** The force is attractive and directed along the straight line joining the objects.

What is the relationship between Eq. 13.1 and the magnitude mg of the gravitational force near Earth's surface? To answer this question, let object 1 in Eq. 13.1 be Earth with mass m_E and let object 2 be an arbitrary object of mass m_o near Earth's surface:

$$F_{Eo}^G = G\frac{m_E m_o}{r_{Eo}^2}. \tag{13.2}$$

For an object a distance h above the ground, the distance between the object and Earth's center is $r_{Eo} = R_E + h$, where R_E is the radius of Earth. Near the ground, $h \ll R_E$, $r_{Eo} \approx R_E$, and Eq. 13.2 becomes

$$F_{Eo}^G = \frac{Gm_E m_o}{r_{Eo}^2} \approx G\frac{m_E m_o}{R_E^2} = m_o\left(\frac{Gm_E}{R_E^2}\right) \quad \text{(near Earth's surface).} \tag{13.3}$$

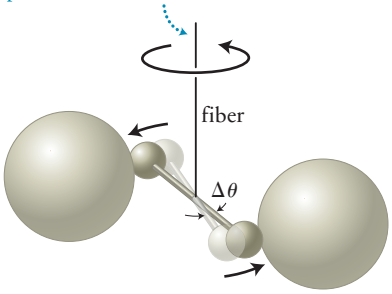
The term in parentheses represents the acceleration of the object due to the force F_{Eo}^G. Knowing from our study of the gravitational force that $F_{Eo}^G = m_o g$, we see that

$$g = \frac{Gm_E}{R_E^2} \quad \text{(near Earth's surface).} \tag{13.4}$$

Because G, m_E, and R_E are all constant, the law of gravity does indeed give a constant acceleration near the surface of Earth.

13.17 For an object released from a height $h \approx R_E$ above the ground, does the acceleration due to gravity decrease, increase, or stay the same as the object falls to Earth?

Figure 13.27 Schematic representation of Cavendish's experimental setup to determine the gravitational constant. The entire setup is enclosed in a box to shield it from air currents.



Working in about 1677, Newton knew the radius of Earth but not its mass, and so he could get only a rough value of G by estimating m_E (as you did in Checkpoint 1.16). The magnitude of G was determined experimentally by Henry Cavendish in 1798. The basic principle of Cavendish's measurement is shown in **Figure 13.27**. Two small lead spheres connected by a light rod are suspended from a long thin fiber. When two large lead spheres are placed near the small spheres, the gravitational attraction between the large and small spheres causes the rod to rotate, thus twisting the fiber. Equilibrium is reached when the torque caused by the gravitational forces is equal in magnitude to the opposing torque caused by the twisted fiber. By calibrating the degree to which known forces twist the fiber, Cavendish was able to measure the gravitational force between the lead spheres. Knowing this force, he then found the value of G by measuring the separation and inertias of the spheres. The value determined with modern instruments is

$$G = 6.6738 \times 10^{-11}\ \text{N}\cdot\text{m}^2/\text{kg}^2. \tag{13.5}$$

QUANTITATIVE TOOLS

Because G is exceedingly small, the force exerted by two 1-kg objects that are 1 m apart is a mere 6.7×10^{-11} N. For comparison, the force of gravity exerted by Earth on a hair is around 10^{-3} N—15 million times stronger! This comparison shows how exceedingly small the gravitational force is. Only when at least one of the two interacting objects has a very large mass (10^{11} kg at the minimum) does the gravitational force become sizeable.

EXAMPLE 13.3 Weighing Earth

Cavendish is said to have "weighed Earth" because his determination of G provided the first value for the planet's mass m_E. Given that the radius of Earth is about 6400 km and given the value of G in Eq. 13.5, determine m_E.

❶ **GETTING STARTED** I am given Earth's radius and G and must use these values to determine Earth's mass. An expression containing G and the acceleration due to gravity g seems a good place to begin.

❷ **DEVISE PLAN** Equation 13.4 gives the acceleration due to gravity near Earth's surface, g, in terms of G, R_E, and m_E. Knowing the values of g, G, and R_E, I can determine m_E.

❸ **EXECUTE PLAN** From Eq. 13.4, I obtain $m_E = gR_E^2/G$,

$$\text{so } m_E = \frac{(9.8 \text{ m/s}^2)(6.4 \times 10^6 \text{ m})^2}{(6.6738 \times 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2)} = 6.0 \times 10^{24} \text{ kg.} \checkmark$$

❹ **EVALUATE RESULT** Table 13.1 gives 5.97×10^{24} kg, which agrees with my answer to within less than 1%. Instead of looking up the value, I can also estimate Earth's mass to check my answer. Remember that we carried out that estimate in Checkpoint 1.16 and obtained 5×10^{24} kg, which is within 20% of my answer.

13.6 Gravitational potential energy

In Section 7.9 we determined the potential energy associated with the force of gravity near Earth's surface. Let us now obtain general expressions for the gravitational potential energy from the law of gravity. To do so, we consider two objects 1 and 2 of masses m_1 and m_2. Suppose the mass of object 1 is much greater than the mass of object 2, $m_1 \gg m_2$, so that if we place object 1 at the origin, we can consider it being fixed there. Now let object 2 move along the x axis from x_i to x_f under the influence of the gravitational attraction between the two (Figure 13.28). What is the work done by the gravitational force exerted by object 1 on object 2?

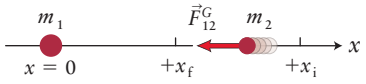


Figure 13.28 An object of mass m_2 moves toward an object of mass $m_1 \gg m_2$ held fixed at the origin.

If we choose the x axis pointing to the right as indicated in Figure 13.28, the x component of the force exerted by object 1 on object 2 when object 2 is at some arbitrary position x is

$$F_{12x}^G = -G\frac{m_1 m_2}{x^2}. \tag{13.6}$$

The x component of the force is negative because the force points in the negative x direction. The gravitational force varies with distance, and so to evaluate the work done by object 1 on object 2, we must evaluate the integral (see Eq. 9.22):

$$W = \int_{x_i}^{x_f} F_x(x)dx. \tag{13.7}$$

Substituting Eq. 13.6 into Eq. 13.7, we obtain

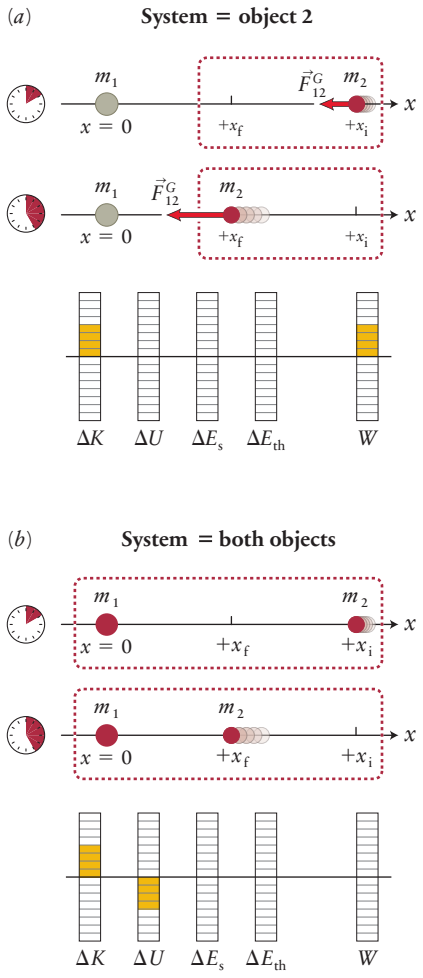
$$W = -Gm_1m_2\int_{x_i}^{x_f}\left(\frac{1}{x^2}\right)dx = Gm_1m_2\left[\frac{1}{x}\right]_{x_i}^{x_f} = Gm_1m_2\left(\frac{1}{x_f} - \frac{1}{x_i}\right). \tag{13.8}$$

This is the work done by the gravitational force exerted by object 1 on the system consisting of object 2 only (Figure 13.29*a* on the next page). Because the force and the displacement point in the same direction, we know that the work must be positive. Indeed, the right side of Eq. 13.8 is positive because $x_f < x_i$.

As object 2 moves from x_i to x_f, it accelerates, and so its kinetic energy increases. Because no other energy associated with it changes, the increase

QUANTITATIVE TOOLS

Figure 13.29 Energy diagrams for the situation shown in Figure 13.28.



in the kinetic energy must be equal to the work done on it by object 1, $\Delta K = W$, as illustrated in the energy diagram in Figure 13.29*a*.

If, instead of object 2 by itself, we now consider the (closed) system of the two interacting objects as a whole, then the increase in the kinetic energy of object 2 is due not to the work done by the external gravitational force exerted by object 1 on it but to a decrease in the gravitational potential energy of the system (Figure 13.29*b* and compare with Figure 9.18). This gravitational potential energy is a measure of the configuration of the system. The change in the gravitational potential energy of the system is thus $\Delta U^G = -\Delta K$. Because the change in kinetic energy does not depend on the choice of system, we see that the change in potential energy of the two-object system is the negative of W in Eq. 13.8:

$$\Delta U^G = Gm_1m_2\left(\frac{1}{x_i} - \frac{1}{x_f}\right). \tag{13.9}$$

It is customary to choose the gravitational potential energy of a system to be zero at infinite separation ($x = \infty$), which is the separation distance at which the gravitational force is zero: $U^G(\infty) = 0$. So, if we let object 2 move in from $x = \infty$ to an arbitrary position x under the influence of the gravitational acceleration due to object 1, Eq. 13.9 yields

$$\Delta U^G = U^G(x) - U^G(\infty) = U^G(x) - 0 = Gm_1m_2\left(0 - \frac{1}{x}\right), \tag{13.10}$$

so the **gravitational potential energy** is

$$U^G(x) = -G\frac{m_1m_2}{x}, \tag{13.11}$$

which is zero at infinite x. As object 2 moves in from infinitely far away, its kinetic energy increases, and so the potential energy of the two-object system decreases. Therefore setting $U^G(\infty) = 0$ at $x = \infty$ causes the potential energy to be negative for all values of $x < \infty$.

EXAMPLE 13.4 Gravitational potential energy near Earth's surface

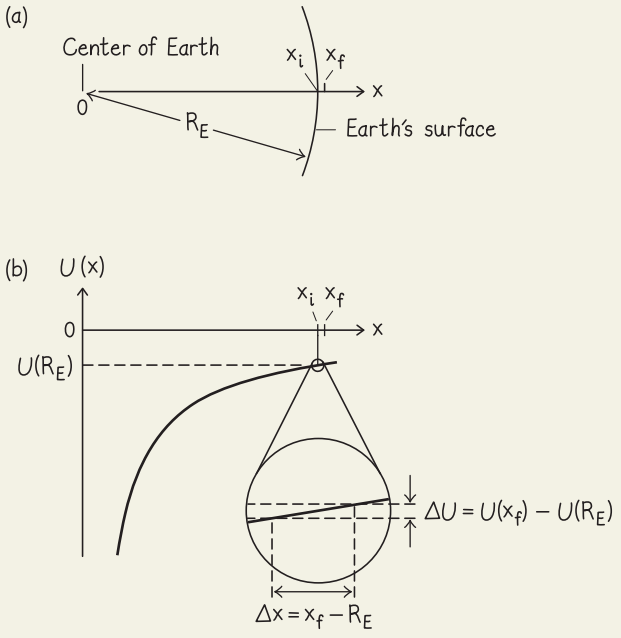
The distance between Earth's surface and an object of mass m is changed by an amount Δx. Show that when $x \approx R_E$ and $\Delta x \ll R_E$, where R_E is the radius of Earth, the general expression for gravitational potential energy, Eq. 13.9, reduces to the expression $\Delta U^G = mg\Delta x$ (Eq. 7.19).

❶ **GETTING STARTED** I begin by choosing an x axis that is perpendicular to Earth's surface and has the origin $x = 0$ at Earth's center. I then mark off a distance $x = R_E$ on this axis (**Figure 13.30***a*). Because I'm told that $x \approx R_E$ and $\Delta x \ll R_E$, I let the initial position x_i be right at Earth's surface, $x = R_E$, and the final position x_f be just to the right of x_i.

To get some feel for the situation, I also graph the gravitational potential energy of Earth from Eq. 13.11, marking the object's initial and final positions (Figure 13.30*b*). Basically, the problem requires me to show that when x_i and x_f lie sufficiently close to each other, ΔU^G is linearly proportional to Δx. Graphically this is already clear from my graph: If I blow up the region around $x_i = R_E$, the $U^G(x)$ curve is linear.

❷ **DEVISE PLAN** If I substitute $x_i = R_E$ and $x_f = R_E + \Delta x$ into Eq. 13.9, I can obtain an expression showing how ΔU^G depends on Δx. I can then approximate the resulting expression for $\Delta x \ll R_E$.

Figure 13.30



QUANTITATIVE TOOLS

3 EXECUTE PLAN After I substitute the masses of Earth and the object, $x_i = R_E$, and $x_f = R_E + \Delta x$, Eq. 13.9 becomes

$$\Delta U^G = Gm_E m\left(\frac{1}{R_E} - \frac{1}{R_E + \Delta x}\right). \qquad (1)$$

I know that $\Delta x \ll R_E$ in this problem, so I can write the second term in the parentheses as

$$\frac{1}{R_E + \Delta x} = \frac{1}{R_E}\left(\frac{1}{1 + \Delta x/R_E}\right) \approx \frac{1}{R_E}\left(1 - \frac{\Delta x}{R_E}\right) \qquad (2)$$

because, for any small number $\epsilon \ll 1$, $1/(1+\epsilon) \approx 1 - \epsilon$ (see Appendix B). Substituting Eq. 2 into Eq. 1, I obtain

$$\Delta U^G = Gm_E m\left[\frac{1}{R_E} - \frac{1}{R_E}\left(1 - \frac{\Delta x}{R_E}\right)\right]$$

$$= Gm_E m\frac{\Delta x}{R_E^2} = m\frac{Gm_E}{R_E^2}\Delta x.$$

When I substitute, from Eq. 13.4, $g = Gm_E/R_E^2$, this expression becomes $\Delta U^G = mg\Delta x$. ✔

4 EVALUATE RESULT This result confirms the conclusion I drew from my sketch: $U^G(x)$ is curved, but for sufficiently small displacements Δx, the curve can be approximated by a straight line. Near Earth's surface the slope of this line is mg.

Because the force of gravity is a central force, we can easily generalize Eq. 13.11 to more than one dimension. Suppose object 2 moves from P to Q in **Figure 13.31*a*** along part of an elliptical orbit. Let us first approximate the displacement of object 2 by two successive displacements: one along the radial direction from P to R and one along a circular arc from R to Q. The work done by the force of gravity on object 2 over the radial displacement is given by Eq. 13.8. Along the circular arc RQ, the magnitude of the force of gravity is constant but always oriented perpendicular to the displacement. Because the work in more than one dimension is given by the scalar product of the force and the force displacement (Eq. 10.35), the work done by the force of gravity on object 2 along the circular arc is zero:

$$W_{R\to Q} = \int_R^Q \vec{F}_{12}^G \cdot d\vec{r} = 0. \qquad (13.12)$$

So the work done by object 1 on object 2 as it moves from P to Q along the path PRQ is equal to the work done along the radial displacement only.

A better approximation of the actual path of object 2 can be obtained by breaking down the motion into small segments and approximating each segment by the sum of radial and circular displacements, as shown in Figure 13.31*b*. Again, the work done on object 2 along each circular displacement is zero. Furthermore, all the radial displacements simply add up to the same radial displacement as in Figure 13.31*a*. So only the change in radial distance matters and the gravitational work done by object 1 on object 2 while object 2 moves from P to Q is

$$W_{P\to Q} = Gm_1m_2\left(\frac{1}{r_Q} - \frac{1}{r_P}\right), \qquad (13.13)$$

where r_P and r_Q are the magnitudes of the position vectors of object 2 at P and Q, respectively. Following the same line of reasoning that led from Eq. 13.8 to Eq. 13.11, we can write for the gravitational potential energy of the system

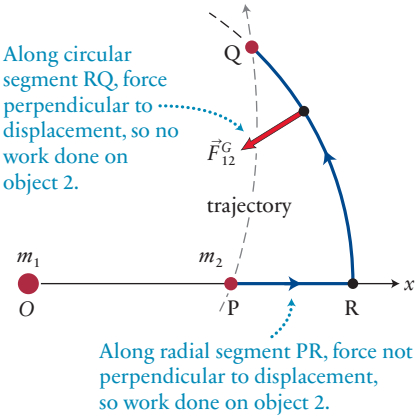
$$U^G(\vec{r}) = -G\frac{m_1m_2}{r} \quad \text{(zero at infinite } r\text{)}, \qquad (13.14)$$

where r is the distance between the two objects.

Note that the result we obtained in Eq. 13.13 holds for *any* path that leads from P to Q. Any arbitrary path can be broken down into small radial and circular segments, and as Figure 13.31*b* shows, the work done by the gravitational force depends only on the position of the endpoints

Figure 13.31 To calculate the work done by object 1 on object 2 while object 2 moves along the dashed path from P to Q, we approximate the path by (*a*) one radial and one circular displacement or (*b*) a number of successive such displacements.

(*a*) Trajectory approximated by one radial segment PR and one circular segment RQ



(*b*) Trajectory approximated by multiple radial and circular segments

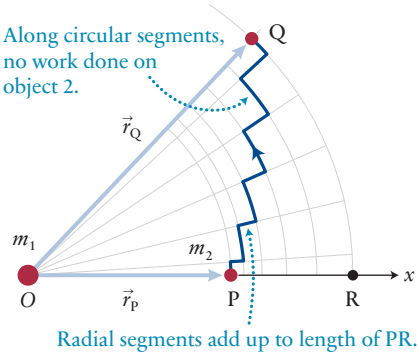
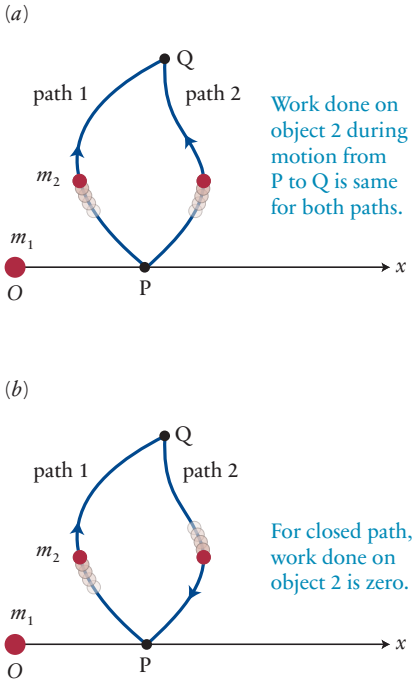


Figure 13.32 The gravitational work done on object 2 depends only on the endpoints.



relative to m_1, not on the path taken. Consider for example paths 1 and 2 from P to Q in **Figure 13.32a**. The work done is the same along both paths:

$$W_{\mathrm{P}\to\mathrm{Q},\ \text{path 1}} = W_{\mathrm{P}\to\mathrm{Q},\ \text{path 2}}. \tag{13.15}$$

This path independence is characteristic of nondissipative interactions (see Section 7.8).

We also know that the work done along a path moving in one direction is the negative of the work done along that path moving in the opposite direction. This can readily be seen in Figure 13.31*b*: Reversing the direction of travel inverts the sign of the displacement (and therefore the sign of the work) along all radial components. So if we let object 2 move from Q to P along path 2 in Figure 13.32*a*, the work done on it is

$$W_{\mathrm{Q}\to\mathrm{P},\ \text{path 2}} = -W_{\mathrm{P}\to\mathrm{Q},\ \text{path 2}}. \tag{13.16}$$

A consequence of Eqs. 13.15 and 13.16 is that if an object moves from P to Q along path 1 and then back to P along path 2, as shown in Figure 13.32*b*, then the work done by the gravitational force on the object is zero:

$$W = W_{\mathrm{P}\to\mathrm{Q},\,\text{path 1}} + W_{\mathrm{Q}\to\mathrm{P},\,\text{path 2}} = W_{\mathrm{P}\to\mathrm{Q},\,\text{path 2}} + (-W_{\mathrm{P}\to\mathrm{Q},\,\text{path 2}}) = 0. \tag{13.17}$$

A path like the one shown in Figure 13.32*b*, in which any object following the path returns to its original position, is called a **closed path.** Because the gravitational potential energy depends only on the position of the endpoints and the endpoints are the same, the change in the potential energy of the two-object system is zero. Consequently the work done by the gravitational force exerted by one object on the other as it moves along a closed path is zero, as Eq. 13.17 shows.

13.18 Near Earth's surface, when $\Delta U^G = mg\Delta x$, it is customary to let the gravitational potential energy be zero at Earth's surface, instead of at infinite separation as we did in Eq. 13.10. Does this choice of zero yield greater or smaller values for the gravitational potential energy? By what amount?

13.7 Celestial mechanics

Now that we have derived an expression for the gravitational potential energy, we can write an expression for the energy of a system of two celestial objects. Consider a system consisting of two objects of which one has a much greater mass than the other ($M \gg m$), so that we can consider the center of mass of the two-object system to be fixed at the center of the larger object (see Section 13.2). Examples of such a system are a planet orbiting a star, a moon orbiting a planet, and a satellite orbiting Earth (**Figure 13.33**). For convenience we'll refer to the object with the greater mass M as the *star* and the object with the smaller mass m as the *satellite,* but the results apply to any pair of celestial objects with $M \gg m$. (Technically, satellites orbit planets, not stars. A planet could have the greater mass M or the smaller mass m, so using "star" throughout this discussion avoids that confusion.) We shall consider the system to be closed and isolated so that neither its energy nor its angular momentum can change:

$$\Delta E = 0 \quad \text{and} \quad \Delta\vec{L} = \vec{0}. \tag{13.18}$$

The energy of the system consists of the gravitational potential energy and the kinetic energy. Because only the satellite is in motion, we have

$$E = K + U^G = \tfrac{1}{2}mv^2 - \frac{GMm}{r}, \tag{13.19}$$

where v is the satellite's speed and r is the distance between satellite and star.

QUANTITATIVE TOOLS

Because the force of gravity is central, the angular momentum of each object is constant. For the satellite we have

$$L = r_{\perp} mv. \tag{13.20}$$

Equations 13.18–13.20 allow us to determine the shape of the satellite's orbit given the values of $\vec{r}$ and $\vec{v}$ at any given instant.

13.19 (*a*) If an object of mass m is released from rest a distance r from a star of mass $M \gg m$ and radius R_s, is the mechanical energy of the star-object system E_{mech} positive, negative, or zero? (*b*) Describe the motion of the object, and determine the maximum and minimum values of its kinetic energy during its motion. (*c*) Does the motion you described in part *b* satisfy conservation of angular momentum about the center of the star? (*d*) Suppose the object is now launched with nonzero velocity in a direction perpendicular to the straight line joining it and the center of the star. Is E_{mech} positive, negative, or zero?

As this checkpoint shows, the mechanical energy E_{mech} of a star-satellite system can be negative because U^G is always negative when we let the gravitational potential energy be zero at infinite separation. To understand the consequences of the energy being negative, look at **Figure 13.34**, which shows the gravitational potential energy U^G as a function of the distance r between satellite and star. As we saw in Section 13.6, U^G has a $1/r$ dependence and approaches zero as r goes to infinity. If the energy E of the system is negative, then there is a value $r = r_{max}$ for which $U^G = E_{mech}$—all of the mechanical energy is in the form of potential energy, and thus the satellite's kinetic energy is zero. Any separation distance r greater than r_{max} would yield a negative kinetic energy because $U^G > E_{mech}$ for $r > r_{max}$. Because the kinetic energy cannot be negative, the satellite's motion for negative values of E_{mech} is restricted to values of $r < r_{max}$ for which $U^G < E_{mech}$. The maximum separation distance r_{max} occurs when $U^G = E_{mech}$:

$$-\frac{GMm}{r_{max}} = E_{mech} \tag{13.21}$$

$$r_{max} = \frac{GMm}{-E_{mech}}. \tag{13.22}$$

In words, we say that for negative E_{mech}, the satellite is *bound* to the star: It cannot escape to infinity because it does not have enough kinetic energy. When the mechanical energy E_{mech} is positive, the gravitational potential energy is *always* smaller than E_{mech}, and so even at infinite separation, where the gravitational potential energy is maximum, there still is a positive amount of kinetic energy. Therefore there is no limit on the separation distance r for $E_{mech} > 0$; in this case, we say the satellite is *unbound* because it has enough energy to "escape" to infinity.

13.20 (*a*) If a satellite of mass m were to reach the position r_{max} given by Eq. 13.22, what would its angular momentum be? (*b*) What would its trajectory have to be to satisfy conservation of angular momentum? (*c*) If its angular momentum is nonzero, can the satellite ever reach r_{max}?

To determine the shape of the satellite's orbit, we must solve Eqs. 13.18–13.20. The solution of these simultaneous equations yields orbits that correspond to the *conic sections* illustrated in **Figure 13.35** on the next page, with the star at the focus of the section. The sections shown in the figure

Figure 13.33 When one object orbits another that has much greater mass, we can consider the center of mass of the system to be fixed at the center of the larger object.

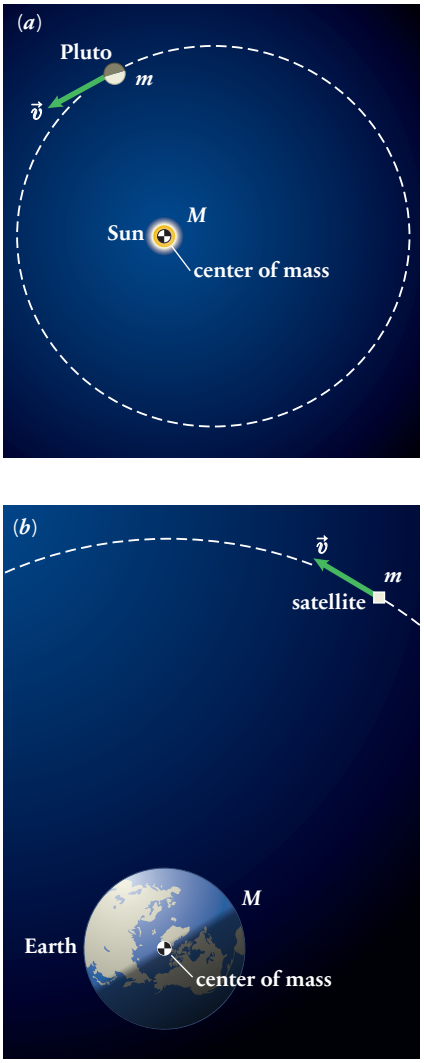
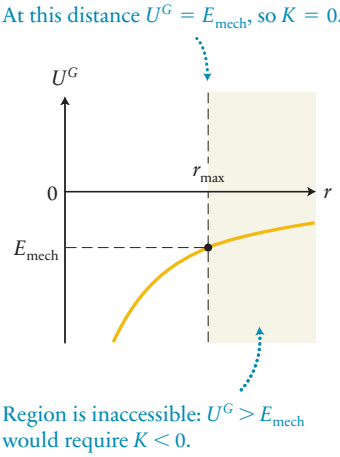


Figure 13.34 When the mechanical energy E_{mech} of a system is negative, the separation distance r between two objects in the system is limited to the region where $U^G < E_{mech}$.



Conic sections

Second-degree equations in two variables of the form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

are called **conic sections** because the curves they represent can be obtained by intersecting a plane and a circular cone (Figure 13.35). These curves can also be written in the form

$$\frac{x^2}{a^2} + \frac{y^2}{a^2(1-e^2)} = 1, \qquad (1)$$

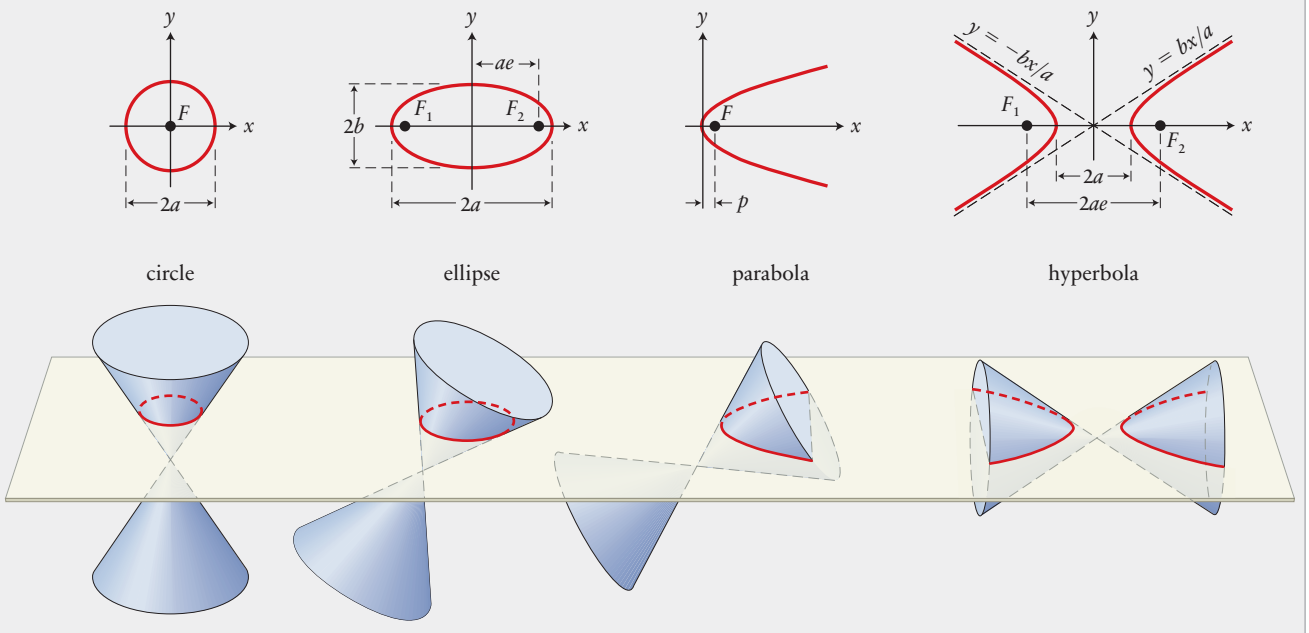
where a is the semimajor axis (see Figure 13.7) of the section and e is the eccentricity.

Ellipse: For $0 \le e < 1$, we obtain an ellipse with a semimajor axis a and a semiminor axis b with $b^2 = a^2(1-e^2)$, where e is the eccentricity of the ellipse. For each point on the ellipse, the sum of the distances to the two foci at $(\pm ae, 0)$ is equal to $2a$. A special type of ellipse is obtained for $e = 0$: The two foci coincide, and the ellipse becomes a **circle** of radius a.

Hyperbola: When $e > 1$, the term that contains y^2 becomes negative, and Eq. 1 yields the two branches of a hyperbola with foci at $(\pm ae, 0)$ and oblique asymptotes $y = \pm(b/a)x$, with $b^2 = a^2(e^2-1)$. For each point on the hyperbola, the difference of the distances to the foci is equal to $2a$.

The limiting case between an ellipse and a hyperbola, which occurs when $e = 1$, yields a **parabola** of the form $y^2 = 4px$. The parabola can be thought of as an elongated ellipse with one focus at $(p, 0)$ and the other at $(\infty, 0)$.

Figure 13.35 Conic sections.



fall into two categories: *ellipses* (including the *circle*, which is a special type of ellipse) and *hyperbolas*. The limiting section between these two categories is the *parabola*. Ellipses are bound and therefore correspond to negative mechanical energy; hyperbolas are unbound and thus correspond to positive mechanical energy. The limiting case between bound and unbound motion occurs for $E_{\text{mech}} = 0$, when the satellite has just enough kinetic energy to approach infinity (with zero kinetic energy). The corresponding orbit is the parabola. Summarizing, we have

$E_{\text{mech}} < 0$: elliptical orbit

$E_{\text{mech}} = 0$: parabolic orbit

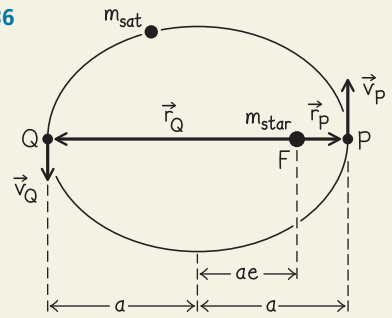
$E_{\text{mech}} > 0$: hyperbolic orbit.

EXAMPLE 13.5 Rocket science

A satellite of mass m_{sat} is in an elliptical orbit around a star of mass $m_{star} \gg m_{sat}$. If the mechanical energy of the star-satellite system is E_{mech} and the magnitude of the angular momentum of the satellite about the star is L, what are the semimajor axis a and the eccentricity e of the satellite's orbit?

1 GETTING STARTED I begin by making a drawing of the system (**Figure 13.36**), placing the star at one focus (F) of the ellipse. I also mark the positions of smallest (P) and and greatest (Q) separation between star and satellite, because at those locations the position vector $\vec{r}$ from star to satellite is perpendicular to the satellite's velocity $\vec{v}$, and therefore it is relatively easy to calculate the angular momentum.

Figure 13.36



2 DEVISE PLAN From my diagram I see that the major axis of the ellipse is equal to the sum of the distances FP and FQ: $2a = r_P + r_Q$. I also see that the distance FP is equal to $a - ae = a(1-e)$, and so $a(1-e) = r_P$. If I know r_P and r_Q, I can determine a and e. I know that the satellite's angular momentum at P must be the same as its angular momentum at Q. I also know that the mechanical energy of the system must be the same at these two locations. Thus I have two equations, $L_P = L_Q = L$ and $E_P = E_Q = E_{mech}$, from which I can determine r_P and r_Q.

3 EXECUTE PLAN The magnitude of the satellite's angular momentum at P is the same as that at Q, so $L = m_{sat} r_P v_P = m_{sat} r_Q v_Q$. Using this expression, I can write for the satellite's kinetic energy at P:

$$K_P = \tfrac{1}{2} m_{sat} v_P^2 = \frac{m_{sat}^2 v_P^2 r_P^2}{2 m_{sat} r_P^2} = \frac{L^2}{2 m_{sat} r_P^2}.$$

The energy of the system at P is then

$$E_{mech} = K_P + U_P^G = \frac{L^2}{2 m_{sat} r_P^2} - \frac{G m_{sat} m_{star}}{r_P}. \quad (1)$$

After writing corresponding expressions for position Q and setting the system energy at P equal to that at Q, I get

$$\frac{L^2}{2 m_{sat} r_P^2} - \frac{G m_{sat} m_{star}}{r_P} = \frac{L^2}{2 m_{sat} r_Q^2} - \frac{G m_{sat} m_{star}}{r_Q}.$$

I now put both terms containing L on the left and combine them over their common denominator, put both terms containing G on the right and combine them over their common denominator, and work through the algebra to get

$$\frac{L^2}{2 m_{sat}} (r_P + r_Q) = G m_{sat} m_{star} r_P r_Q.$$

Substituting $r_P + r_Q = 2a$, $r_P = a(1-e)$, and $r_Q = 2a - r_P = a(1+e)$, I obtain, after some algebra,

$$L^2 = G m_{sat}^2 m_{star} a (1 - e^2). \quad (2)$$

Substituting this expression for L^2 into Eq. 1 and solving for a give

$$a = -\frac{G m_{sat} m_{star}}{2 E_{mech}}. \checkmark$$

Substituting this result for a into Eq. 2 and solving for e, I finally obtain

$$e = \sqrt{1 + \frac{2 E_{mech} L^2}{G^2 m_{sat}^3 m_{star}^2}}. \checkmark$$

4 EVALUATE RESULT The orbit being elliptical means that E_{mech} is negative, and so a is positive and e is smaller than 1, as they should be for an ellipse. When $E_{mech} = 0$, the orbit should turn into a parabola and the satellite can go to infinity. Indeed, as I let E_{mech} approach zero, a approaches infinity and $e = 1$, and the ellipse turns into a parabola (see the box "Conic sections" on page 448).



13.21 Consider a planet of mass m moving at constant speed v in a circular orbit of radius R under the influence of the gravitational attraction of a star of mass M. (*a*) What is the planet's kinetic energy, in terms of m, M, G, and R? (*b*) What is the energy of the star-planet system? (*c*) Using the expression for the eccentricity obtained in Example 13.5, obtain an expression for E_{mech} for the circular orbit. (*d*) Compare your answers to parts *b* and *c*.

The results obtained in Example 13.5 show that the length $2a$ of the major axis of the ellipse depends on only the mechanical energy E_{mech} of the system. Because E_{mech} is negative, the greater E_{mech}, the closer it is to zero and the greater a.

Figure 13.37 Orbits of a satellite of mass m launched in different directions with fixed speed and a fixed distance from an object of mass M. Because the mechanical energy E_{mech} of the system is always the same, the length $2a$ of the orbit's major axis (dashed line) is constant. As the angle between $\vec{v}_i$ and $\vec{r}_i$ is decreased, the angular momentum decreases and the eccentricity of the orbit increases.

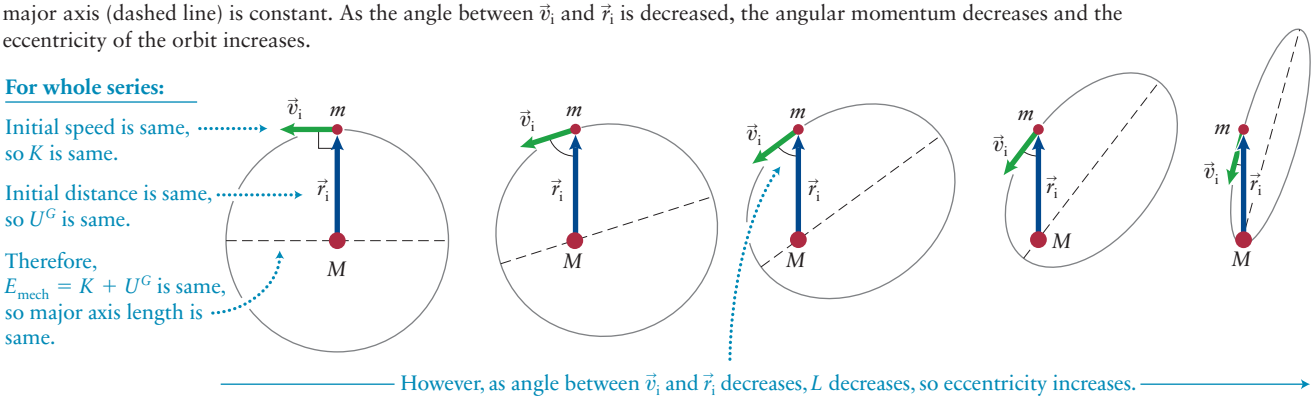
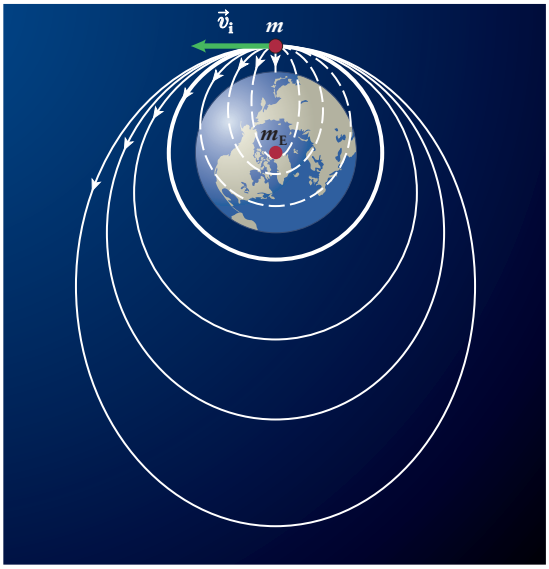
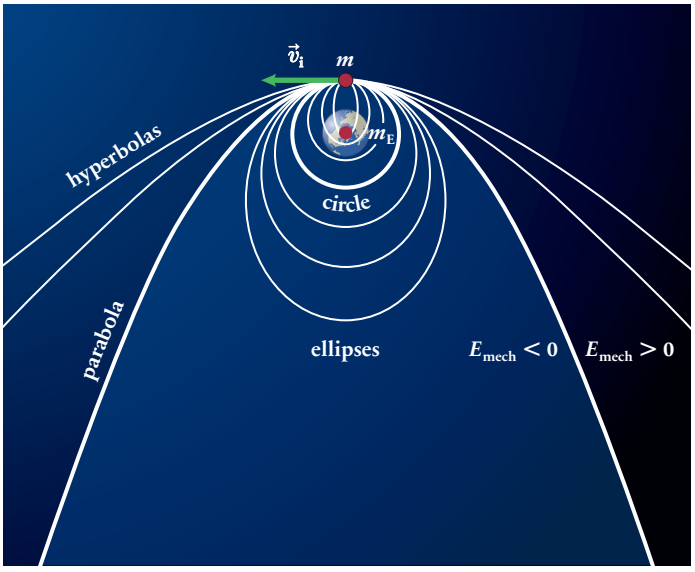


Figure 13.37 shows five orbits with the same fixed energy E_{mech} but different values of L. The energy $E_{mech} = U^G + K$ is fixed because the satellite is launched at a fixed initial speed (so K is fixed) from a point P that is a fixed distance r_i from the center of the larger object of mass M (so U^G is fixed). Because E_{mech} is fixed, the length $2a$ of the major axis is the same for all orbits. As the direction of $\vec{v}_i$ is changed (without changing the speed v_i and therefore E_{mech} and $2a$), the magnitude of the angular momentum changes, affecting the eccentricity of the orbit, which depends on both L and E_{mech}. For a fixed mechanical energy E_{mech} (that is to say, for a fixed value of the major axis), e increases as L decreases. So, as the angle between $\vec{v}_i$ and $\vec{r}_i$ gets smaller, the ellipse becomes more elongated. In all cases, the object of mass M is at one of the foci of the ellipse.

Figure 13.38 shows the orbits of an object launched multiple times from a fixed location that is a distance r_i from Earth's center. All launches are in the same direction, perpendicular to the straight line connecting the object to Earth's center, but the launch speed is increased for each successive launch. Because the initial gravitational potential energy U^G is always the same, the value of the mechanical energy E_{mech} is determined by the initial kinetic energy: $E_{mech} = U^G + K$.

Figure 13.38 Orbits of an object launched at different speeds from a fixed location that is a fixed distance r_i from Earth's center.





If the object is released from rest, $K = 0$ and so $E_{\text{mech}} = U^G$. As you saw in Checkpoint 13.19, the object falls in a straight line toward the surface of Earth. When the object is given a small initial speed, however, the orbit is an ellipse with Earth's center at the focus farther from the launch point. For small v_i, the orbit intersects Earth and the object crashes. As v_i is increased, the ellipse becomes rounder and longer, and for sufficiently great v_i, the object makes it around the planet.

As we continue to increase v_i, we reach a point when v_i has just the right magnitude to make v_i^2/r_i equal to the centripetal acceleration caused by Earth's gravitational force, and the object goes in a circular orbit. For greater v_i, the orbit is again elliptical, but Earth's center now coincides with the nearer focus of the ellipse.

As we continue to increase v_i, we reach a point where the mechanical energy E_{mech} of the Earth-object system is no longer negative and the orbit is unbound. This happens when

$$E_{\text{mech}} = \tfrac{1}{2}mv_{\text{esc}}^2 - \frac{Gmm_E}{r_i} = 0, \tag{13.23}$$

where v_{esc} is the object's *escape speed*, m is the mass of the launched object, m_E is Earth's mass, and r_i is the distance from the launch point to Earth's center. An object launched at the escape speed and moving without any further propulsion travels out to infinity along a parabolic path. At launch speeds greater than v_{esc} the orbit corresponds to one branch of a hyperbola.



13.22 (*a*) Determine an expression for the escape speed at Earth's surface. (*b*) What is the value of this escape speed? (*c*) Does it matter in which direction an object is fired at the escape speed?

13.8 Gravitational force exerted by a sphere

Central to Newton's law of gravity is the statement that a solid sphere exerts a gravitational force on an object outside the sphere as if all its mass were concentrated at the center of the sphere. To prove this statement, we'll calculate the force exerted on a particle of mass m by a thin spherical shell of mass M and radius R; the distance between the particle and the center of the shell is r. Once we have a result for the shell, we can obtain the result for a solid sphere by treating the solid sphere as a large number of concentric spherical shells, like the layers of an onion.

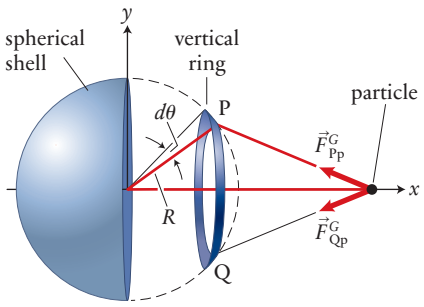
Consider first a vertical ring-shaped piece of the shell—the portion contained between the angles θ and $\theta + d\theta$ in **Figure 13.39a**. The force exerted by a small segment of the ring near P is $\vec{F}^G_{\text{Pp}}$; the force by a small segment near Q is $\vec{F}^G_{\text{Qp}}$. Only the components of these forces along the x axis contribute to the vector sum of the forces exerted on the particle. The same cancellation occurs for every pair of diametrically opposite small segments that make up the entire ring. To calculate the vector sum of the forces exerted on the particle, we therefore need to consider only the x components. The x component of the force exerted by a small segment of the ring of mass dm near P on the particle is

$$dF^G_{\text{Pp}x} = -G\frac{m\,dm}{s^2}\cos\alpha, \tag{13.24}$$

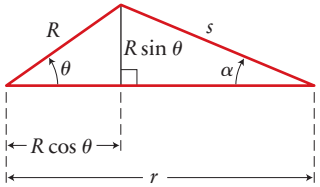
where α is the angle between the force and the straight line from the particle to the center of the shell and s is the distance from the particle to P. The x components of the forces exerted by all segments of the ring have the

Figure 13.39 To calculate the gravitational force exerted by a spherical shell of mass M on an object of mass m, the shell is divided into vertical rings.

(*a*)



(*b*) Geometry of red triangle in part *a*.



same magnitude. The vector sum of the forces exerted by the whole ring is thus equal to the sum of the contributions from all segments of the ring:

$$dF_{\mathrm{rp}x}^{G} = -G\frac{m\,dM}{s^2}\cos\alpha, \tag{13.25}$$

where dM is the mass of the entire ring (the sum of the masses of all segments). The width of the ring being equal to the arc length $R\,d\theta$, the area of the ring is (length) $\times$ (width) $= (2\pi R\sin\theta)(R\,d\theta) = 2\pi R^2\sin\theta\,d\theta$. If the mass M of the entire shell is uniformly distributed over the shell, then the mass per unit area is $M/(4\pi R^2)$, so

$$dM = 2\pi R^2\sin\theta\,d\theta\frac{M}{4\pi R^2} = \frac{M}{2}\sin\theta\,d\theta \tag{13.26}$$

or, combining Eqs. 13.25 and 13.26,

$$dF_{\mathrm{rp}x}^{G} = -\frac{GMm}{2}\frac{\cos\alpha\sin\theta\,d\theta}{s^2}. \tag{13.27}$$

Equation 13.27 gives the force exerted by one vertical ring of our spherical shell on the particle. Now we must integrate this result over the entire shell. Before we can carry out this integration, we must reduce the variables α, s, and θ, which are all related, to one variable. From Figure 13.39*b* we see that

$$\cos\alpha = \frac{r - R\cos\theta}{s}, \tag{13.28}$$

and from the law of cosines (see Appendix B), we have

$$s^2 = R^2 + r^2 - 2rR\cos\theta. \tag{13.29}$$

Differentiating the left side of this equation with respect to s yields $2s$. For the right side we get

$$\begin{aligned}\frac{d}{ds}(R^2 + r^2 - 2rR\cos\theta) &= \frac{d}{d\theta}(R^2 + r^2 - 2rR\cos\theta)\frac{d\theta}{ds}\\ &= 2rR\sin\theta\frac{d\theta}{ds}.\end{aligned} \tag{13.30}$$

Equating the derivatives on the two sides, we get

$$2s = 2rR\sin\theta\frac{d\theta}{ds} \tag{13.31}$$

$$\sin\theta\,d\theta = \frac{s\,ds}{rR}. \tag{13.32}$$

Substituting Eq. 13.32 and $R\cos\theta$ from Eq. 13.29 into Eq. 13.28 and then into Eq. 13.27 yields

$$dF_{\mathrm{rp}x}^{G} = -\frac{GMm}{4r^2R}\frac{(r^2 + s^2 - R^2)ds}{s^2} = -\frac{GMm}{4r^2R}\left(\frac{r^2 - R^2}{s^2} + 1\right)ds. \tag{13.33}$$

QUANTITATIVE TOOLS

To obtain the force exerted by the entire shell, we must integrate this expression from $s = r - R$ (at $\theta = 0$) to $s = r + R$ (at $\theta = \pi$):

$$F^G_{\text{sp}x} = \int_{r-R}^{r+R} dF^G_{\text{rp}x} = -\frac{GMm}{4r^2R}\int_{r-R}^{r+R}\left(\frac{r^2 - R^2}{s^2} + 1\right)ds. \tag{13.34}$$

Working out the integral (see Checkpoint 13.23), we get

$$\int_{r-R}^{r+R}\left(\frac{r^2 - R^2}{s^2} + 1\right)ds = 4R \tag{13.35}$$

so

$$F^G_{\text{sp}x} = -G\frac{Mm}{r^2} \quad \text{(shell)}, \tag{13.36}$$

where the minus sign indicates that the force points toward the center of the shell. So we see that the shell exerts a force as if its entire mass M were concentrated at the center.

We can now easily extend this result to a solid sphere by considering many concentric shells. For each shell, we obtain an expression of the form given by Eq. 13.36; adding all terms (one for each shell) then yields

$$F^G_{\text{sp}} = G\frac{mM_{\text{sphere}}}{r^2} \quad \text{(sphere)} \tag{13.37}$$

for the magnitude of the gravitational force exerted by a solid sphere of mass M_{sphere} on a particle of mass m placed a distance r from the sphere's center. Equation 13.37 shows that the solid sphere exerts a force as if its entire mass M were concentrated at the center.



13.23 (*a*) Work out the integral in Eq. 13.35. (*b*) Show that Eq. 13.33 still holds if the particle of mass m is inside the shell. (*c*) What are integration limits with this particle inside the shell? What is the value of the integral in Eq. 13.35 for these limits?

Chapter Summary

Universal gravity and the principle of equivalence (Sections 13.1, 13.3–13.5, 13.8)

Concepts The force of gravity is an attraction between all objects that have **mass,** which is a measure of the quantity of material in an object. Mass is the property of an object that determines the strength of its gravitational interaction with other objects.

At everyday speeds, the mass of an object is equal to its inertia.

According to the **principle of equivalence,** one cannot distinguish locally between the effects of a constant gravitational acceleration of magnitude g and the effects of an acceleration of the reference frame of magnitude $a = g$.

Quantitative Tools According to Newton's **law of gravity,** if two objects that have masses m_1 and m_2 are a distance r apart, the magnitude of the gravitational force that each object exerts on the other is

$$F_{12}^G = G\frac{m_1 m_2}{r_{12}^2}, \tag{13.1}$$

where $G = 6.6738 \times 10^{-11}\ \mathrm{N \cdot m^2/kg^2}$ is the **gravitational constant.**

The acceleration due to gravity near Earth's surface is

$$g = \frac{Gm_E}{R_E^2}, \tag{13.4}$$

where m_E is Earth's mass and R_E is its radius.

The gravitational force F_{sp}^G exerted by a uniform solid sphere on a particle of mass m is the same as if all the matter in the sphere were concentrated at its center:

$$F_{sp}^G = G\frac{mM_{sphere}}{r^2}, \tag{13.37}$$

where M_{sphere} is the mass of the sphere and r is the distance from its center to the particle.

Angular momentum and gravitational potential energy (Sections 13.2, 13.6)

Concepts The force of gravity is a **central force,** meaning that its line of action always lies along the line connecting the interacting objects.

In an isolated system of two objects that interact through a central force, each object has a constant angular momentum about the center of mass of the system.

Quantitative Tools The **gravitational potential energy** $U^G(x)$ of a system made up of two objects of masses m_1 and m_2 that are a distance x apart is

$$U^G(x) = -G\frac{m_1 m_2}{x}, \tag{13.11}$$

where we take $U^G(x)$ to be zero at $x = \infty$.

Celestial mechanics (Sections 13.2, 13.7)

Concepts **Kepler's first law of planetary motion:** All planets move in elliptical orbits with the Sun at one focus. This law is a consequence of the fact that the magnitude of the gravitational force has a $1/r^2$ dependence.

Kepler's second law: The straight line from any planet to the Sun sweeps out equal areas in equal time intervals. This law reflects the fact that the angular momentum of a planet is constant because the gravitational force is a central force.

Kepler's third law: The squares of the periods T of the planets are proportional to the cubes of the semimajor axes a of their orbits around the Sun ($T^2 \propto a^3$). This law is a consequence of the fact that the magnitude of the gravitational force has a $1/r^2$ dependence.

Quantitative Tools The energy of a closed ($\Delta E = 0$), isolated ($\Delta L = 0$) system made up of a small object of mass m orbiting a large object of mass $M \gg m$ is

$$E = K + U^G = \tfrac{1}{2}mv^2 - G\frac{Mm}{r}, \tag{13.19}$$

where v is the speed of the small object, r is the distance between the objects, and the large object does not move.

The magnitude of angular momentum of the small object is

$$L = r_\perp mv. \tag{13.20}$$

The value of E in Eq. 13.19 determines the orbit's shape:

$E_{mech} < 0$: elliptical (bound) orbit
$E_{mech} = 0$: parabolic orbit
$E_{mech} > 0$: hyperbolic (unbound) orbit.

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

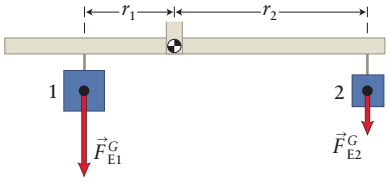
Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

[It is convenient for some problems to note that $G = 6.67 \times 10^{-11}\ \text{N} \cdot \text{m}^2/\text{kg}^2 = 6.67 \times 10^{-11}\ \text{m}^3/(\text{kg} \cdot \text{s}^2)$.]

13.1 Universal gravity

1. If the matter that makes up a planet is distributed uniformly so that the planet has a fixed, uniform density, how does the magnitude of the acceleration due to gravity at the planet surface depend on the planet radius? •
2. In Figure P13.2, suppose the mass of object 1 is three times the mass of object 2. (*a*) At Earth's surface, what ratio r_1/r_2 is needed to have the rod stay horizontal? (*b*) What is this ratio on the Moon, where the magnitude of the gravitational force is only one-sixth of the magnitude of Earth's gravitational force? •

Figure P13.2

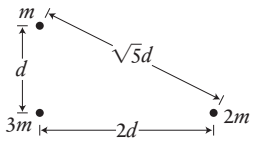


3. Suppose you're making a scale model of the solar system showing the eight planets (not including the dwarf planet Pluto) and having each planet at its farthest position from the Sun. If you use a marble of radius 10 mm for Mercury, how far from the Sun must you place your outermost planet? •
4. Four identical objects are placed at the four corners of a square, far from any star or planet. (*a*) Choose one object and draw to scale the vectors that represent the gravitational forces exerted on it by the other three objects. (*b*) Draw a vector that represents the vector sum of the forces exerted on your chosen object. ••
5. Two basketballs, each of mass 0.60 kg and radius 0.12 m, are placed on a floor so that they touch each other. Two golf balls, each of mass 0.045 kg and radius 22 mm, are placed on a table so that they touch each other. What is the ratio of the gravitational force exerted by one basketball on the other to that exerted by one golf ball on the other? ••

13.2 Gravity and angular momentum

6. An astronaut on the International Space Station gently releases a satellite that has a mass much smaller than the mass of the station. Describe the motion of the satellite after release. •
7. An isolated object of mass m can be split into two parts of masses m_1 and m_2. Suppose the centers of these parts are then separated by a distance r. What ratio of masses m_1/m_2 would produce the largest gravitational force on each part? ••
8. Particles of mass m, $2m$, and $3m$ are arranged as shown in Figure P13.8, far from any other objects. These three particles interact only gravitationally, so that each particle experiences a vector sum of forces due to the other two. Call these $\vec{F}_m$, $\vec{F}_{2m}$, and $\vec{F}_{3m}$. (*a*) Show that the lines of action of $\vec{F}_m$, $\vec{F}_{2m}$, and $\vec{F}_{3m}$ intersect at a common point. (*b*) Is this point of intersection the center of mass of the system? (*c*) Is the analysis of the motion of this system straightforward? ••

Figure P13.8

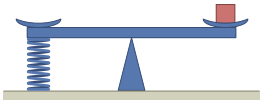


9. Angular momentum can be represented as a vector product and interpreted in terms of the rate of change of an area with respect to time. Give a similar interpretation of another vector product: torque. (Hint: Consider derivatives.) •••

13.3 Weight

10. Suppose the spring/balance device in Figure P13.10 is adjusted so that the spring is slightly stretched and the balance arm is level when the device is in an elevator at rest relative to the ground and a 1.0-kg object is placed as shown. If the elevator accelerates upward, does the balance arm tilt clockwise, tilt counterclockwise, or remain level? •

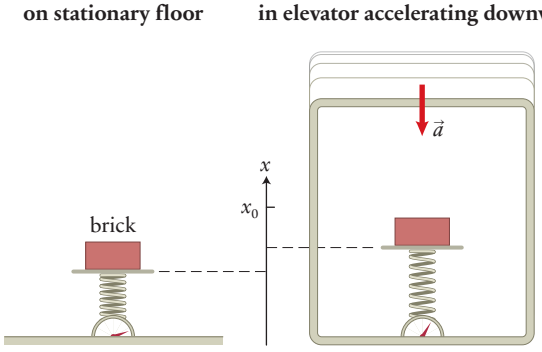
Figure P13.10



11. A certain metallic structure weighs 2450 N on earth. The structure is sent to Mars, which has a diameter of 6.78×10^6 m and mass of 6.42×10^{23} kg. What is the weight of the structure on the surface of the Martian planet? ••
12. A balance is purposefully prepared in an unbalanced condition, with too much mass on one side. It is carefully held with two hands, one on the base of the balance and one on the heavier end, so that the balance arm is level. Then both hands release the system from rest at the same instant. While it is in free fall, what happens to the balance arm? ••

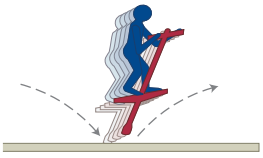
13. (*a*) For what downward acceleration does the spring scale in Figure P13.13 read zero? (*b*) What would the answer in part *a* be if the experiment were done on the Moon? (*c*) How would your conclusion about the spring forces change if the brick were hung from the spring scale rather than supported by it? (Ignore the weight of the spring.) ••

Figure P13.13



14. You ride your pogo stick (Figure P13.14) across a diving board and into the local swimming pool. Your pogo stick is a special model equipped with a sensor that records the force measured by the spring. Describe the values of force recorded during your ride, using the gravitational force exerted on you by Earth as a benchmark. ••

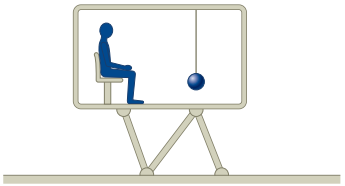
Figure P13.14



13.4 Principle of equivalence

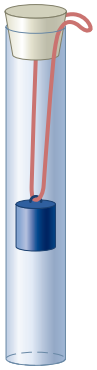
15. A person in an elevator lets go of his briefcase but observed that it moves towards the ceiling. How is the elevator moving? ••
16. How should the simulator in Figure P13.16 be tilted to simulate a left turn in a bobsled? Assume that the bobsled run has banked turns and that the sled exerts no tangential force on the ice. ••

Figure P13.16



17. A vertical accelerometer for measuring accelerations on a roller coaster can be constructed using a piece of clear plastic pipe, a cork, a rubber band, and a metal bob (Figure P13.17). Where the bob hangs down normally is where you make a mark for an acceleration of *g*. Without accelerating the pipe, explain how you can calibrate the accelerometer as you assemble it for accelerations of 2*g*, 3*g*, and so on. (Extra bobs are available.) ••

Figure P13.17



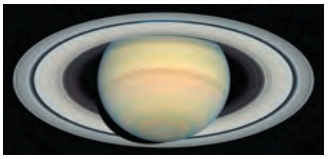
18. Consider two satellites simultaneously launched 180 km apart at the equator, each placed in a circular orbit that passes above the North Pole and its launch point. (*a*) Describe their relative motion with respect to Earth and with respect to each other. (*b*) Does this motion imply that each satellite exerts a force on the other? ••
19. Let us explore whether there is any way to distinguish acceleration due to rotation from acceleration due to a gravitational force. Imagine a deep bowl with a small volume of milk in it. (*a*) What happens to the milk when you spin the bowl about a perpendicular axis that runs through the center of the bowl bottom? (*b*) What is the shape of an object that exerts a gravitational force that could make the milk behave in a similar fashion? •••

13.5 Gravitational constant

20. Consider two speres in a system—Sphere 1 has mass M and radius $2R$, while the second Sphere has mass $2M$ and radius R. If the two spheres touch each other, find the magnitude of their gravitational force of attraction as a function of mass M and radius R. •
21. Two spherical celestial bodies, one of mass 3.5×10^{13} kg and the other of mass 2.4×10^{18} kg, experience a gravitational force between them with a magnitude of 4.0×10^{7} N. How far apart are the two spheres? ••
22. A 2.0×10^{9}-kg sphere is at the origin, and a second 5.0×10^{9}-kg sphere is on the x-axis at $x = 8.0$ km. Find the magnitude and direction of the net gravitational attraction force that the two spheres would exert on a third sphere of mass 9.1×10^{3} kg placed at $x = 2.0$ km. ••
23. The mass of a hypothetical planet is 1/80 that of Earth and its radius is 1/4 that of Earth. What is the ratio of the acceleration due to gravity on the planet to that on Earth? ••
24. Gravity is the weakest of the fundamental interactions (see Section 7.6). (*a*) This being so, what makes it such an important force on Earth? (*b*) What makes it the dominant force in galaxies? ••
25. A neutron star has about two times the mass of our Sun but has collapsed to a radius of 10 km. What is the acceleration due to gravity on the surface of this star in terms of the free-fall acceleration at Earth's surface? ••
26. Which pulls harder on the Moon: Earth or the Sun? ••
27. How far above the surface of Earth do you have to go before the acceleration due to gravity drops by 0.10%, 1.0% and 10%? ••
28. (*a*) As a spacecraft travels along a straight line from Earth to the Moon, at what distance from Earth does the force of gravity exerted by Earth on the coasting spacecraft cancel the force of gravity exerted by the Moon on the spacecraft? (*b*) What do the passengers in the spacecraft notice, if anything, before and after passing this location? ••

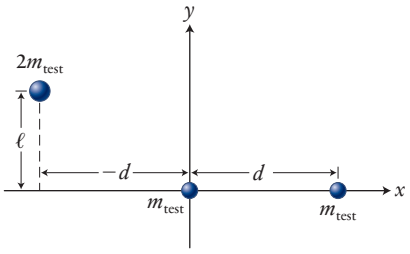
29. From the picture of Saturn in Figure P13.29 and known values, estimate the orbital period of particles in Saturn's rings. ●●

Figure P13.29



30. The Sun exerts a gravitational force on the Earth, keeping it in its orbit. (*a*) As per Newton's third law, what is the reaction to this force? (*b*) Suppose another planet has the same mass as the Earth but is located twice as far from the Sun. Describe the magnitude of the attraction force. ●
31. The acceleration of gravity on the surface of the Moon is about 1/6 of Earth's *g*. If the radius of the moon is 1,737.4 km, approximate the mass of the Moon from the given information. ●●
32. At what height above the surface of Pluto is the acceleration due to gravity half its surface value? ●●
33. Three spheres are arranged on a line along the *x*-axis. Sphere $m_1 = 5.00$ kg and is at the origin while sphere m_2 is at $x = +35.0$ m. What must be the mass of the second sphere if the net gravitational attraction force experience by a third sphere of mass 2.00 kg at $x = +20.0$ m is zero? ●●

Figure P13.33



34. Using *g* for the acceleration due to gravity is valid as long as the height *h* above Earth's surface is small. Derive an expression for a more accurate value of the acceleration due to gravity as a quadratic function of *h*. ●●●
35. Calculate the acceleration due to gravity *inside* Earth as a function of the radial distance *r* from the planet's center. (Hint: Imagine that a mine shaft has been drilled from the surface to Earth's center and an object of mass *m* has been dropped down the shaft to some radial position $r < R_E$ from the center, where R_E is the radius of Earth. What is the vector sum of the forces exerted on the object by all of Earth's mass that lies at any distance $d > r$ from the center? $d < r$ from the center?) ●●●

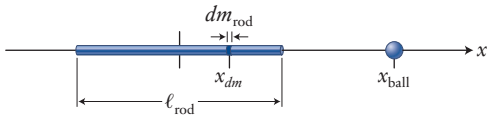
13.6 Gravitational potential energy

36. The Earth has a mass of 5.972×10^{24} kg, while the Moon has a mass of 7.348×10^{22} kg. Calculate the gravitational potential energy of the Earth-Moon system when the distance between the two bodies is 3.844×10^5 km. ●
37. The asteroid known as Toro has a radius of about 5.0 km and a mass of 2.0×10^{15} kg. Could a 70-kg person standing on its surface jump free of Toro? ●
38. (*a*) How much gravitational potential energy does a system comprising a 150-kg object and the Moon have if the object is one Moon radius above the ground? (*b*) How far should this object be from the moon for the gravitational potential become zero? ●●
39. An object 1 of mass m_1 is separated by some distance *d* from an object 2 of mass $2m_1$. An object 3 of mass m_3 is to be placed between them. If the potential energy of the three-object system is to be a maximum (closest to zero), should object 3 be placed closer to object 1, closer to object 2, or halfway between them? ●●
40. Two particles, each of mass *m*, are initially at rest very far apart. Obtain an expression for their relative speed of approach at any instant as a function of their separation distance *d* if the only interaction is their gravitational attraction to each other. ●●
41. Consider an alternate universe in which the magnitude of the attractive force of gravity exerted by Earth on a meteor of mass m_m approaching Earth is given by $F = Cm_m m_E/r^3$, where *C* is some positive constant and *r* is the center-to-center distance between Earth and the meteor. (*a*) How much work is done by the gravitational force on the meteor as it falls from very far away (infinite distance) to some height *h* above Earth's surface? (*b*) If the meteor is moving very slowly when it is very far away, how fast is it moving when it gets to that height? ●●
42. Derive expressions for the orbital speed and energy for a satellite of mass m_s traveling in a circular orbit of radius *a* around a planet of mass $m_p \gg m_s$. ●●
43. What maximum height above the surface of Earth does an object attain if it is launched upward at 4.0 km/s from the surface? ●●
44. (*a*) Derive an expression for the energy needed to launch an object from the surface of Earth to a height *h* above the surface. (*b*) Ignoring Earth's rotation, how much energy is needed to get the same object into orbit at height *h*? ●●
45. A spacecraft propelled by a solar sail is pushed directly away from the Sun with a force of magnitude C/r^2, where *C* is a constant and *r* is the spacecraft-Sun radial distance. The craft has a mass of 5.0×10^4 kg and starts from rest at a distance $r_i = 1.0 \times 10^8$ km from the Sun. What must the numerical value and units of *C* be in order for the craft to have a speed of 0.10*c* when it is very far from the Sun? (Hint: $c = 3.00 \times 10^8$ m/s.) ●●
46. In 1865, Jules Verne wrote a story in which three men went to the Moon by means of a shell shot from a large cannon sunk in the ground. (*a*) What muzzle speed must the cannon have in order for the shell to reach the Moon? (*b*) If the acceleration was constant throughout the muzzle, how long would the muzzle need to be in order to achieve this speed without exposing the men to an acceleration greater than 6*g*? (*c*) Argue whether this launch method is practical or not. ●●
47. The center-to-center distance between a 200-g lead sphere and an 800-g lead sphere is 0.120 m. A 1.00-g object is placed 0.0800 m from the center of the 800-g sphere along the line joining the centers of the two spheres. (*a*) Ignoring all sources of gravitational force except the two spheres, calculate the gravitational force exerted on the object. (*b*) Determine the gravitational

potential energy per gram at the position of the object. (*c*) How much work is needed to keep the object from getting closer than 0.0400 m from to the 800-g sphere? ●●

48. Derive an expression for the gravitational potential energy of a system consisting of Earth and a brick of mass m placed at Earth's center. Take the potential energy for the system with the brick placed at infinity to be zero. ●●●
49. Is approximating the gravitational potential energy near the surface of the Sun as $mg_S\Delta x$ accurate over a smaller or larger range of values than approximating the gravitational potential energy near the surface of Earth as $mg_E\Delta x$? ●●●
50. A uniform rod of mass m_{rod} and length ℓ_{rod} lies along an x axis far from any stars or planets, with the center of the rod at the origin (Figure P13.50). A ball of mass m_{ball} is located at position x_{ball} on the axis. (*a*) Write an expression for the gravitational potential energy of the system made up of the ball and a small element dm_{rod} of the rod located at position x_{dm}. (*b*) Integrate your expression over the length of the rod to determine the potential energy of the system for $x_{ball} > \ell_{rod}/2$. (*c*) It can be shown that for such a system the gravitational force exerted by the rod on an object located a distance x from the origin is $F_x = -dU/dx$. Using this relationship, compute the force exerted by the rod on the ball at any distance x from the rod center. ●●●

Figure P13.50

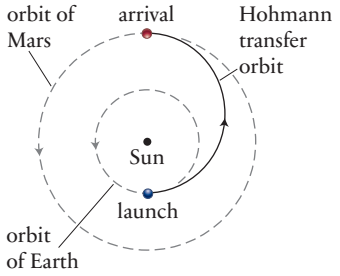


13.7 Celestial mechanics

51. The parabolic orbit of any comet around the Sun might be described as a collision between the two objects. Would it be better described as an elastic collision or an inelastic collision? ●
52. Assume a particle is located at the surface of the Sun. (*a*) If the particle has mass m, how fast must it be moving away from the Sun's center of mass to escape the gravitational influence of the Sun? (*b*) At what speed must a particle of mass $2m$ move in order to escape the Sun's gravitational influence? ●
53. What is the speed of a space probe when it is very far from Earth if it was launched from the surface of Earth at twice its escape speed? ●
54. Which trip takes more rocket fuel: from Earth to the Moon or from the Moon to Earth? ●●
55. A satellite moves at speed v in very low orbit about a moon that has no atmosphere. You launch a projectile vertically from the same moon's surface at the same speed v. To what height does it rise? ●●
56. A satellite in an elliptical orbit around Earth has a speed of 8032 m/s when it is at perigee, the position in the orbit closest to Earth. At this position, the satellite is 112 km above Earth's surface. How far above the ground is the satellite when it is at apogee, the position in its orbit farthest from Earth? ●●
57. A certain comet comes in on a parabolic approach around the Sun. Its perihelion is about the same distance from the Sun as the orbit of Mercury. How much faster than Mercury (expressed as a ratio) is the comet traveling at perihelion? ●●
58. Show that for circular orbits of two objects about their center of mass, $E = \frac{1}{2}U = -K$. This is a special case of the more general mathematical statement known as the *virial theorem*. ●●
59. An attractive central force exerted on a particle of mass m causes the particle to travel in a bound orbit. The orbit is elliptical with semimajor axis a, and the force has magnitude Cm/r^2, where C is a constant and r is the orbit radius. If the particle's speed is $v = \sqrt{C/2a}$ when it is farthest from the source of the force, what is its speed when it is nearest the force source? ●●
60. A particle is leaving the Moon in a direction that is radially outward from both the Moon and Earth. What speed must it have to escape the Moon's gravitational influence? ●●
61. A meteoroid passes through a position in space where its speed is very small relative to Earth's and it is at a perpendicular distance of 19 Earth radii above Earth's surface. The meteoroid is moving in such a way that Earth captures it. What is the speed of the meteoroid when it is one Earth radius above the ground? ●●
62. A space probe can be either launched at escape speed from Earth or transported to a "parking orbit" above Earth and then launched. (*a*) If the parking orbit is 180 km above Earth's surface, what is the escape speed from the orbit? (*b*) What launch speed is required to have the probe escape Earth if it is launched from the ground? ●●
63. Two identical stars, each of mass 3.0×10^{30} kg, revolve about a common center of mass that is 1.0×10^{11} m from the center of either star. (*a*) What is the rotational speed of the stars? (*b*) If a meteoroid passes through the center of mass, perpendicular to the orbital plane of the stars, how fast must the meteoroid be going if it is to escape the gravitational attraction of the stars? ●●
64. Kepler's third law says that the square of the period of any planet is proportional to the cube of its mean distance from the Sun. Show that this law holds for elliptical orbits if one uses the semimajor axis as the mean distance. (Said another way, orbits that have the same energy but different angular momentum have the same period.) ●●●
65. Two celestial bodies of masses m_1 and m_2 are orbiting their center of mass at a center-to-center distance of d. Assume each body travels in a circular orbit about the center of mass of the system, and derive the general Newtonian form of Kepler's third law: $T^2 = 4\pi^2 d^3/G(m_1 + m_2)$. ●●●
66. To travel between Earth and any other planet requires consideration of such things as expenditure of fuel energy and travel time. To simplify the calculations, one chooses a path such that the position in Earth's orbit where the launch occurs and the position in the other planet's orbit when the spacecraft arrives define a line that passes through the Sun, as shown for an Earth-Mars transfer in Figure P13.66. The path the spacecraft travels is an ellipse that has the Sun at one focus. Such a path is called a Hohmann transfer orbit, and the major axis of the ellipse, $2a$, is the sum of the radii of the Earth's and the other planet's orbits around the Sun. Take $m_S = 1.99 \times 10^{30}$ kg, $m_M = 6.42 \times 10^{23}$ kg, $m_E = 5.97 \times 10^{24}$ kg, $a_E = 1.50 \times 10^{11}$ m, and $a_M = 2.28 \times 10^{11}$ m, and assume the planets have circular

orbits. (*a*) What is the energy of the system comprising the Sun and a 1000-kg space probe in a Hohmann transfer orbit? (*b*) What is the speed of the probe in this orbit, as a function of r, the probe's radial distance from the Sun? (*c*) What is the probe's speed relative to the Sun as the probe enters the transfer orbit? (*d*) What is its speed relative to the Sun as it approaches Mars? (*e*) Given that the orbital speed of Earth is about 2.98×10^4 m/s, how much additional speed does the probe need to begin the transfer orbit? (*f*) What is the required launch speed from Earth's surface for a probe traveling to Mars in a Hohmann transfer orbit? •••

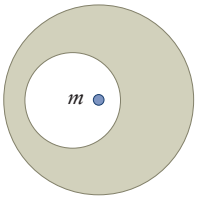
Figure P13.66



13.8 Gravitational force exerted by a sphere

67. A small disk of mass m is at the center of a larger disk that has an off-center hole drilled in it (Figure P13.67). What is the direction of the gravitational force exerted by the larger disk on the small one? •

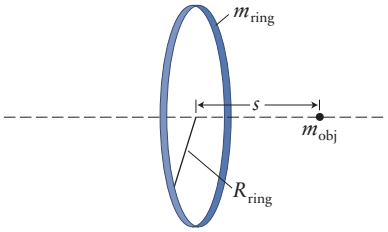
Figure P13.67



68. Is Eq. 13.37 for the gravitational force exerted by a solid sphere of mass m_s on some object of mass m_o when the two are a radial distance r apart, $F_{so}^G = Gm_o m_s/r^2$, valid for a sphere whose mass is not distributed uniformly? •

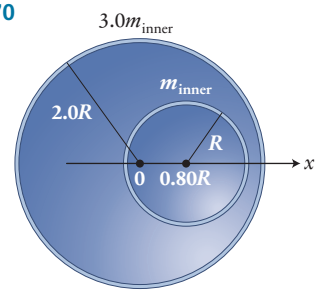
69. A uniform ring of mass m_{ring} and radius R_{ring} is shown in Figure P13.69. A small object of mass m_{obj} sits a separation distance s from the ring on the line that is perpendicular to the plane of the ring and passes through the ring's center. Obtain an expression (in terms of the given variables and universal constants) for the magnitude of the gravitational force exerted by the ring on the object. ••

Figure P13.69



70. Two uniform spherical shells are located such that their centers are along the x axis in Figure P13.70. The inner shell has mass m_{inner} and radius R, and its center is at $x = 0.80R$. The outer shell has mass $3.0m_{inner}$ and radius $2.0R$, and its center is at $x = 0$. What is the magnitude of the vector sum of the gravitational forces exerted by these two shells on an object of mass m_{obj} placed (*a*) at $x = 3.0R$, (*b*) at $x = 1.9R$, and (*c*) at $x = 0.90R$? ••

Figure P13.70



71. A thin disk of radius R has mass m_{disk} uniformly distributed over its area. (*a*) Derive an expression for the magnitude of the gravitational force this disk exerts on a particle of mass m_{part} located a distance y directly above the center of the disk. (*b*) If $y \gg R$, show that this expression reduces to that for two particles. •••

Additional Problems

72. In a system of N particles, how many terms are there in the expression for the gravitational potential energy of the system? •

73. You equip a rocket with enough fuel to reach escape speed from Earth. If you plan to launch the rocket from just above ground level, does it matter whether you launch it vertically, at an angle, or horizontally? •

74. What is the value of the acceleration g_{Venus} due to gravity at the surface of Venus? •

75. Is there any position in the elliptical orbit of a satellite where the tangential component of the acceleration is greater than the component perpendicular to the tangential component? If so, what conditions on the orbit must there be for such a position to exist? ••

76. The *gravitational vector field* $\vec{g}(\vec{r})$ is defined as the gravitational force exerted on a small object located at a particular position $\vec{r}$ divided by that object's mass. (*a*) Sketch the gravitational vector field produced by Earth at a sampling of positions around Earth. Make sure your vector lengths indicate the correct relative sizes. (*b*) Express the gravitational force exerted by Earth on an object of mass m_{obj} located a distance r_{obj} ($> r_E$) from Earth's center in terms of Earth's gravitational vector field. ••

77. Some writers of fiction have placed human beings on asteroids or other "planetoids" whose radius, in the illustrations accompanying the stories, seems not too much greater than the height of the people themselves! What would the density of such a celestial body have to be for the human to feel a gravitational attraction numerically equal to the attraction we feel to Earth? ••

78. You wish to put a satellite in orbit around Earth. (*a*) Which takes more energy per kilogram of satellite: launching the satellite from the ground to a height of 1600 km above the ground or placing the satellite into orbit once it has reached that altitude? (*b*) Repeat for an orbit 3200 km above the ground. (*c*) Repeat for an orbit 4800 km above the ground. ●●

79. Unable to sleep, you crawl out of bed one evening and stare out your window at the night sky, finding Jupiter and sitting back to contemplate its grandeur. Remembering your physics, you imagine yourself at some height above the apparent surface of Jupiter, experiencing a gravitational force that has a magnitude equal to the magnitude of the gravitational force you experience at the surface of Earth. ●●● CR

80. After seeing an old movie, you become concerned about a large asteroid that could be gravitationally attracted to Earth from very far away. You realize that long ago such an asteroid's speed relative to Earth's might have been very small, and thus you are able to calculate the relative speed it has at any separation distance from Earth. In order to get a baseline, you assume that the only object exerting a gravitational force on the asteroid is Earth, and you ignore the speed of Earth in its path around the Sun. What worries you is that observers might not sight the asteroid until it passes the imaginary line defining the Moon's orbit around Earth, and how little time that would give humanity to react. ●●● CR

81. You are working on a project to plot the course of a spy satellite. The satellite has a polar orbit (which means it passes over both of Earth's poles during each orbit), and it is outfitted with a camera carrying a wide-angle lens that can "see" strips of land up to 2500 km wide. The enemy is believed to have chosen an elevation of orbit that just guarantees complete coverage of Earth each day by these surveillance strips. You cannot avoid or block the satellite unless you know where it is. ●●● CR

82. While helping an astronomy professor, you discover a binary star system in which the two stars are in circular orbits about the system's center of mass. From their color and brightness, you determine that each star has the same mass as our Sun. The orbital period of the pair is 24.3 days, based on the oscillation of brightness observed as one star occludes (hides) the other. From this information you are able to ascertain the distance between the stars. ●●● CR

83. Scientists still are not entirely sure whether the universe is open (which means it will keep expanding forever) or closed (which means it will eventually fall back in on itself in a "Big Crunch"). What determines whether it is open or closed is basically the same thing we calculate to determine escape speeds. Assume that the universe has a uniform density ρ. Then the mass contained in some spherical volume of space of radius R is $\frac{4}{3}\pi R^3\rho$. Model a galaxy located at the edge of this spherical volume as a particle and consider its energy. According to Hubble's law, the speed of this galaxy is HR, where $H = 2.2 \times 10^{-18}\ \text{s}^{-1}$ is called the Hubble constant. (*a*) What is the maximum density of the universe such that it does not close? (*b*) The current estimate of the density of the universe is $10^{-26}\ \text{kg/m}^3$, including about 25% *dark matter,* which is matter we cannot observe directly, and 70% *dark energy*, another constituent that is not directly observable. Both dark matter and dark energy contribute to gravity, though, and so are included in the density estimate. Compare this value with your result and discuss what difference it would make if there were no dark matter in the universe. ●●●



14

Special Relativity

14.1 Time measurements
14.2 *Simultaneous* is a relative term
14.3 Space-time
14.4 Matter and energy

14.5 Time dilation
14.6 Length contraction
14.7 Conservation of momentum
14.8 Conservation of energy

CONCEPTS

QUANTITATIVE TOOLS

Time is the basic fabric of life because living means *changing* and change—a concept central to all we have discussed up to now—cannot exist without time. Time is so basic that we take for granted that we all agree on what it is. If, for instance, you and I arrange to meet here tomorrow at three o'clock and we make sure our watches agree now, we expect to see each other here tomorrow when our watches read three o'clock.

As we see in this chapter, however, our commonsense experience of time fails to explain certain experimental observations. In 1905 Albert Einstein realized these shortcomings and revised the basic notions of space and time with his *theory of special relativity*. The basic principles of this theory are the principle of relativity, which we discussed in Chapter 6, and the experimental fact that the measured value of the speed of light is independent of the motion of the source relative to the observer making the measurement. The revised notions of space and time require us to adjust our understanding of inertia, momentum, and energy. However, the conservation principles—the unifying principles of this text and the most fundamental laws of physics—continue to hold, as required by the principle of relativity.

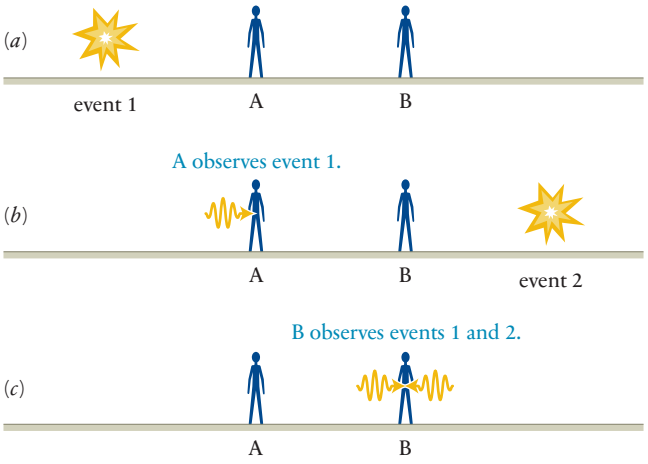
14.1 Time measurements

Since the advent of clocks, the measurement of time has been so straightforward that it hardly requires much thought. Just as spatial coordinates determine *where* something happens, time tells us *when* it happens. Let's define an **event** as something that happens at a specific location at a specific instant. Take, for example, your act of reading this sentence. If you and I agree on a coordinate grid (for example, longitude, latitude, and altitude) and a time zone, you can specify the location and time of this event in such a way that it is uniquely and precisely defined. You are reading this sentence, say, in Helena, Montana (46.595°N, 112.030°W, alt. 1138 m), at 9:33 p.m. today.

Let's concentrate for now on the time part of describing events. By referring to the watch on your wrist, you can measure the instant at which any event near you occurs. There are two reasons you can count on this procedure to work. First, it is possible to construct a device—a clock—that measures equal, calibrated time intervals (say, seconds). Clocks serve as a "ruler" for time, just as an ordinary ruler helps us measure space. Second, we can *synchronize* clocks; that is, we can make sure two clocks run at the same rate and their readings agree.

Keeping time helps us sort out the sequence of events and establish a possible causal relationship between events (see Section 1.4). If you are at the location of two successive events, it's straightforward to determine which event happened first and what the time interval between them is. How, though, can you determine the sequence for two events that occur at different locations?

Figure 14.1 Two observers with synchronized clocks observe two events.



We need to be able to detect some sort of signal (such as sound or light) to observe an event, and all such signals travel at some finite speed. Of all signals known, light travels the fastest and therefore is the most convenient for timing events.

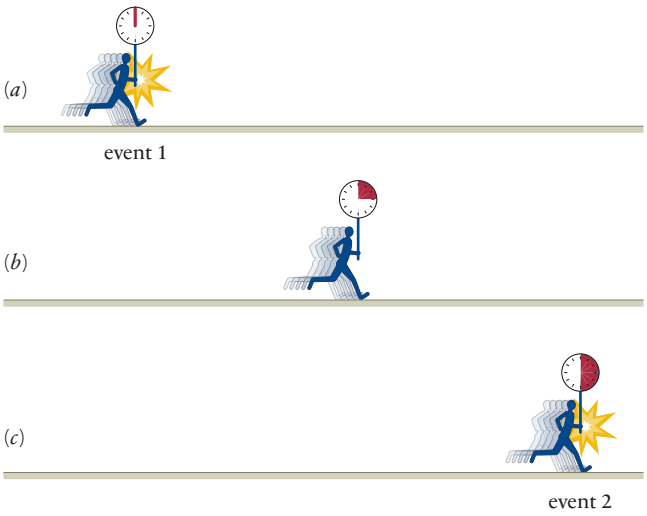
Figure 14.1 shows two events—the emission of light signals at two locations—and their observation by two observers using synchronized clocks. The light signals are of short duration, like a flash, so that we can consider each emission to take place at a definite instant. The signals travel at the speed of light (3.00×10^8 m/s in vacuum or air) and carry the message that the event has occurred. Observer A is closer to event 1 than to event 2; observer B is closer to event 2 than to event 1. In Figure 14.1*a*, event 1 occurs but the signal has not yet reached either observer. In Figure 14.1*b*, the signal from event 1 reaches observer A and event 2 occurs. Note that the signal from event 1 has not yet reached observer B and that the signal from event 2 has reached neither observer. A little later, the signals from both events reach observer B (Figure 14.1*c*) but the signal from event 2 has not yet reached observer A.

14.1 In Figure 14.1, based on the time of arrival of the light signals, which event is observed first according to (*a*) observer A and (*b*) observer B?

As this checkpoint shows, if you go by the instants at which the signals reach the two observers, the observers do not agree on the sequence of the two events. Observer A sees the events as occurring at two different instants, but observer B sees them as occurring at the same instant. The instant of *observation*, however, is not the instant at which the light signals *are emitted*. For events that occur around you in daily life, the difference between the instant each event happens and the instant you observe it is negligible because light travels so fast. Light emitted by the Sun, however, takes about 8 minutes to reach us. Thus, if a solar flare occurs at the surface of the Sun, that event takes place

CONCEPTS

Figure 14.2 One way to measure the time interval between two events is to travel from one event to the other, using a single clock to record the instant at which each occurs.



8 minutes before we observe it. Light emitted by the Andromeda Galaxy takes 2 million years to reach Earth. If you look at the Andromeda Galaxy on a dark, moonless night, you see the light it emitted 2 million years ago. A lot could have happened to that galaxy in those 2 million years—for all you know, the entire galaxy might not be there anymore!

Even though we do not deal with these astronomical distances here in our discussion of time measurements, we still need to take into account the *signal delay* between event and observation. This signal delay is the time interval it takes for the signal to travel from the location of the event to the location of the observer. For nearby events, the signal delay is negligible. For events separated by an appreciable distance, we can avoid signal-delay complications in several ways. One straightforward, if not always practical, way is shown in **Figure 14.2**: An observer travels from the location of event 1 to the location of event 2 and records the two events on the same clock. In the observer's reference frame, the two events take place *at the same location* (the origin if you imagine axes attached to the observer) but at different instants. To avoid the complications of a noninertial reference frame (see Chapter 6), the observer should travel at constant velocity. The time interval obtained this way is called the **proper time interval:***

> **The time interval between any two events that occur at the same position in a particular reference frame is called the proper time interval between the two events.**

Not all pairs of events are separated by a proper time interval: If an observer does not have enough time to travel from the first event to the second event to observe both events at his location, there is no proper time interval for those events.

*From the French *propre* meaning *own*, as in one's "own time."

EXAMPLE 14.1 Proper time interval

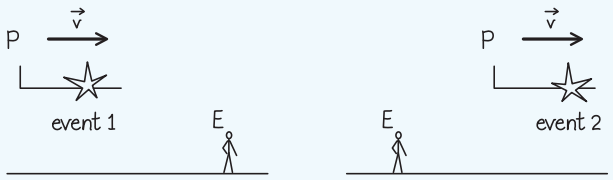
You are flying in a plane, and a friend is on the ground. Which of you measures the proper time interval between (*a*) the beginning and end of a brief in-flight video program and (*b*) two successive signals from a warning light on a tall building along your flight path?

1 GETTING STARTED The problem involves two observers (my friend and I) and two pairs of related events. My friend is in the Earth reference frame (E); I am in the reference frame of the plane (P). To simplify the discussion, I denote the beginning and end of the movie as events 1 and 2, respectively, and the two successive flashes of the warning light as events 3 and 4.

2 DEVISE PLAN To see who measures the proper time interval, I should make a sketch for each pair of events, showing the positions of the E and P reference frames. If either observer sees either pair of events as occurring at the same location, that observer measures the proper time interval.

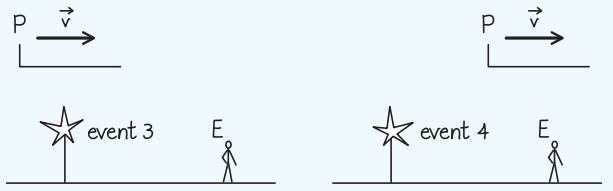
3 EXECUTE PLAN (*a*) **Figure 14.3** shows that the beginning (event 1) and end (event 2) of the video program take place at the same location in reference frame P. Because that is my reference frame, I measure the proper time interval between these events. Figure 14.3 also shows that to my friend, standing at one location in the Earth reference frame, the two events happen at different locations (event 1 high and to his left, event 2 high and to his right). Thus he cannot measure the proper time interval. ✔

Figure 14.3



(*b*) **Figure 14.4** shows that the two flashes take place at the same place in reference frame E. Because this is my friend's reference frame, he measures the proper time interval between the two flashes. As for me, I see the first flash as coming from right below the plane but the second flash as coming from behind the plane. Two different locations from my point of view mean that I cannot measure the proper time interval between the flashes. ✔

Figure 14.4



4 EVALUATE RESULT The video "happens" in the plane, so the plane is the reference frame in which to measure the proper time interval—the screen on which the video is projected is at rest in the plane. Likewise, the warning light is at rest in the Earth reference frame. My answers make sense.

CONCEPTS

Because measuring a proper time interval is not always practical, another way to measure the time interval between events that occur in two locations is to place observers with identical clocks near each event, so that each observer can record the nearby event according to his clock (**Figure 14.5a**). Once the events are over, the two observers can communicate their clock readings and determine which event occurred first.

To synchronize the clocks, the observers can use the procedure illustrated in Figure 14.5*b*. An event sends out light signals to both observers, and each observer records the instant at which the signal reaches his clock. They then measure the distance from each clock to the location of the event and correct for the signal delay. If the clocks are synchronized, the observers obtain the same instant for the event after they subtract the signal delay from their recorded clock reading. If they don't obtain the same instant, they can use this procedure to adjust their clocks.

Because we don't always know where events are going to occur, it makes sense to construct a reference frame that has identical clocks arranged in a grid at regular fixed spacings, as shown in **Figure 14.6**. We can synchronize the clocks by sending a signal from the clock at the origin (the *reference clock*) to all the other clocks and correcting each clock's reading for the signal delay. We can then define the instant in that reference frame at which any event occurs as the reading on the clock whose position coincides with that event. From now on, when we say "measuring the instant at which an event occurs in a given reference frame," we shall understand this to mean that we are using this procedure and that *an observer* in this reference frame is the observer stationed at the clock nearest the event.

Now that we have established a procedure for determining the instant at which an event occurs in a certain reference frame, we can determine the order of events in that reference frame. Two events that occur at different positions in the reference frame are **simultaneous** when the readings of the two clocks at the two locations of the events are identical. When the events are not simultaneous, the clock readings at the locations of the events allow us to determine the sequence of the events.

Figure 14.5 Synchronized clocks can be used to measure the time interval between events that occur in different locations. (*a*) Each observer records the instant at which a nearby event occurs. (*b*) To synchronize their clocks, the two observers record the instant at which a single event occurs.

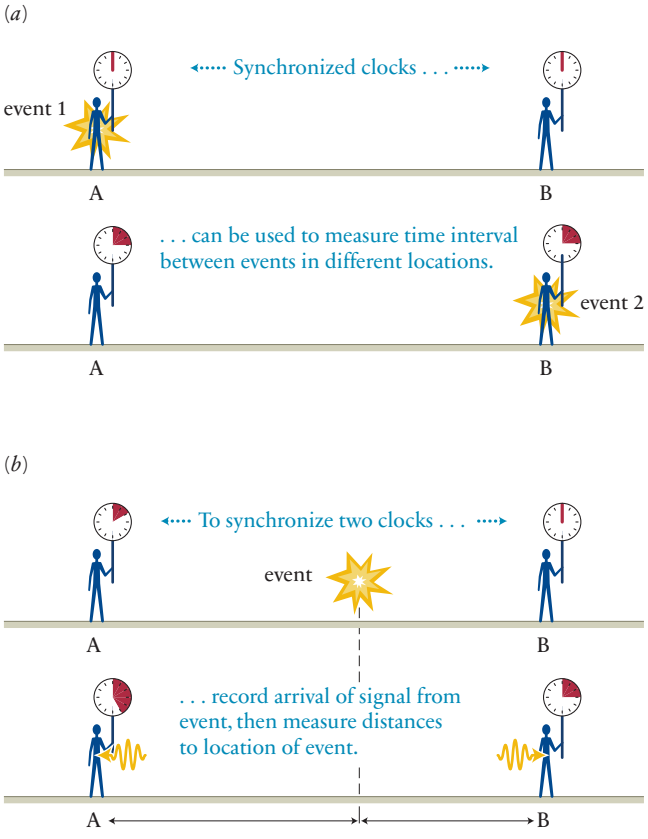
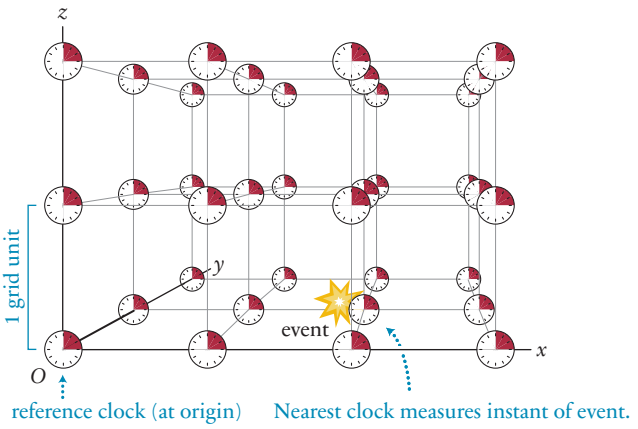


Figure 14.6 The scaffold of time: a reference frame made up of a grid of synchronized clocks. An event can occur anywhere in the grid, and the instant at which the event occurs is the instant shown on the clock nearest the event.



EXERCISE 14.2 Synchronizing clocks

Suppose you want to synchronize the network of clocks in Figure 14.6 by emitting a single light pulse at the origin. This signal triggers each clock to start running from some initial setting. If it takes the signal 1 s to travel along one grid unit and if the reference clock is set initially to 12 o'clock, what should the initial settings be for each clock (*a*) along the x axis and (*b*) along the diagonal in the $y = 0$ plane?

SOLUTION (*a*) Along the x axis, the signal arrives at the first clock to the right of the reference clock 1 s after the emission, when the reference clock reads 12:00:01, and so I set the first clock to 12:00:01. The next clock along the x axis starts running when the signal triggers it at 12:00:02, and so I set each clock on the x axis 1 s ahead of the clock immediately to its left. ✔

CONCEPTS

(*b*) The light must travel $\sqrt{2}$ grid units to reach the first clock along the diagonal. We must therefore set that clock ahead of the reference clock by $\sqrt{2}$ s = 1.4 s. The next clock along the diagonal is set ahead by $2\sqrt{2}$ s, and so on. ✔

14.2 Suppose an observer is standing at the origin in Figure 14.6 and all the clocks are synchronized. (*a*) Which of the following is true? The observer sees that (*i*) all the clocks read the same time, (*ii*) nearby clocks display later times than more distant clocks, (*iii*) nearby clocks display earlier times than more distant clocks. (*b*) Does the observer see all the clocks running at the same rate?

14.2 *Simultaneous* is a relative term

Suppose we determine that two events—for example, the emission of light signals by two sources—are simultaneous in a certain reference frame. Are these events also simultaneous in another reference frame that is in motion relative to the first? To help answer this question, we use a *simultaneity detector*: a device that emits a blue light when it receives two signals at the same instant.

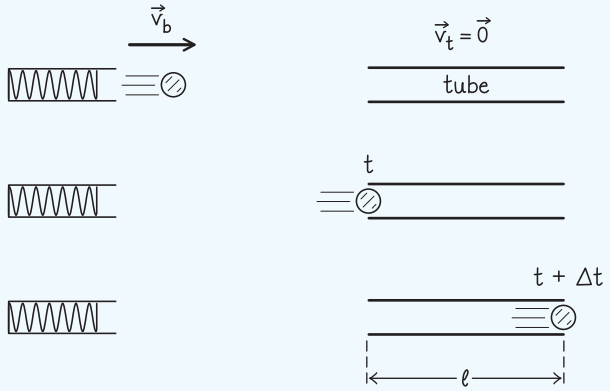
When we position the detector midway between the positions where the two events take place, as illustrated in **Figure 14.7**, the time interval it takes to travel to the detector is the same for both signals. This means that, if the two signals reach the detector at the same instant and the blue light comes on, the events were simultaneous in the reference frame of the detector (Figure 14.7*a*). If the events occur at different instants, the two signals do not arrive at the detector at the same instant and the blue light does not come on (Figure 14.7*b*). The blue light thus tells us whether or not the two emission events were simultaneous *in the reference frame of the detector*. One advantage of this device is that we do not need to be in the reference frame of the detector to determine whether the two events are simultaneous in its reference frame. As long as the simultaneity detector is placed midway between the two events, the blue light tells us whether the events are simultaneous in the detector's reference frame.

Up to now we have discussed only measurements made by an observer (or detector) at rest with respect to the sources of the signals. What happens to the transmission of signals when the source and observer are moving relative to each other? To find out, let's start by reviewing the relativity of speeds for everyday objects that move much more slowly than light.

EXAMPLE 14.3 Fast ball

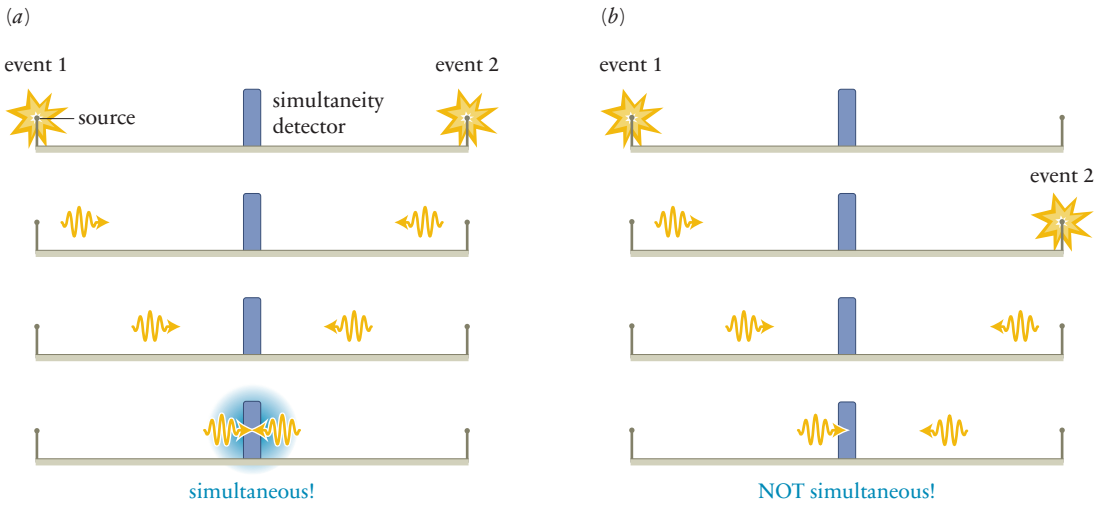
Consider a spring-loaded device that ejects a ball at some unknown everyday speed v_b. To measure this speed, you measure the time interval it takes the ball to pass through a stationary tube of known length ℓ (**Figure 14.8**). Now suppose the tube is moving at speed v_t toward the ejection device. What speed do you obtain by dividing the length ℓ of the tube by the time interval Δt it takes the ball to pass through the tube? Express your answer in terms of v_b and v_t.

Figure 14.8 Example 14.3.



(Continued)

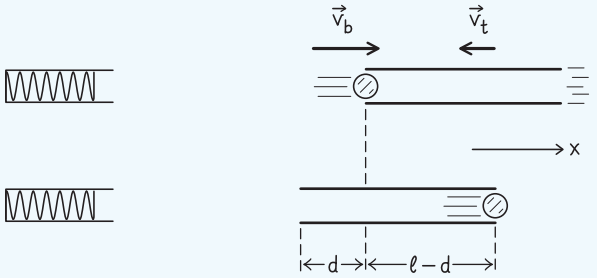
Figure 14.7 A detector located midway between two light sources can detect whether signals are emitted (*a*) simultaneously or (*b*) not simultaneously.



CONCEPTS

1 GETTING STARTED I begin by drawing the ball at the instant it enters and the instant it exits the moving tube (**Figure 14.9**). I choose my positive x axis to be in the direction of the ball's motion and use d to denote the distance the tube moves while the ball passes through. The distance the ball travels as it passes through the tube is thus $\ell - d$.

Figure 14.9



2 DEVISE PLAN From my sketch I see that during the time interval Δt, two things happen: The tube advances a distance d, and the ball travels a distance $\ell - d$. Because speed equals distance traveled divided by the time interval needed to travel this distance, I can relate these two distances to the speeds v_b and v_t. This gives me two equations, from which I can solve for the two unknowns d and Δt. I also note that dividing ℓ by Δt gives me the speed v_{Mb}, which is the speed of the ball measured by an observer M moving along with the tube.

3 EXECUTE PLAN Because the tube advances a distance d in the time interval Δt, I have $v_t = d/\Delta t$. Likewise, for the ball: $v_b = (\ell - d)/\Delta t = \ell/\Delta t - d/\Delta t$. The first term equals the speed v_{Mb} I am asked to determine. The second term is the speed of the tube v_t, and so $v_b = v_{Mb} - v_t$. Rearranging terms, I get $v_{Mb} = v_b + v_t$. The speed of the tube is added to that of the ball. ✔

4 EVALUATE RESULT Using the subscript cancellation rule (Eq. 6.14), I can write $\vec{v}_{Mb} = \vec{v}_{ME} + \vec{v}_{Eb}$. With my choice of positive x axis to the right in Figure 14.9, I get $v_{Mbx} = v_{MEx} + v_{Ebx}$. Because $\vec{v}_{ME}$ is the velocity of Earth relative to an observer moving with the tube, I have $v_{MEx} = v_t$. Likewise $\vec{v}_{Eb}$ is the velocity of the ball relative to Earth, and so $v_{Ebx} = v_b$. Therefore, $v_{Mbx} = v_{MEx} + v_{Ebx} = v_t + v_b$, which confirms the result I obtained.

14.3 In Example 14.3 what is the time interval Δt it takes the ball to pass through the moving tube if $v_b = v_t = v$? Express your answer in terms of ℓ and v.

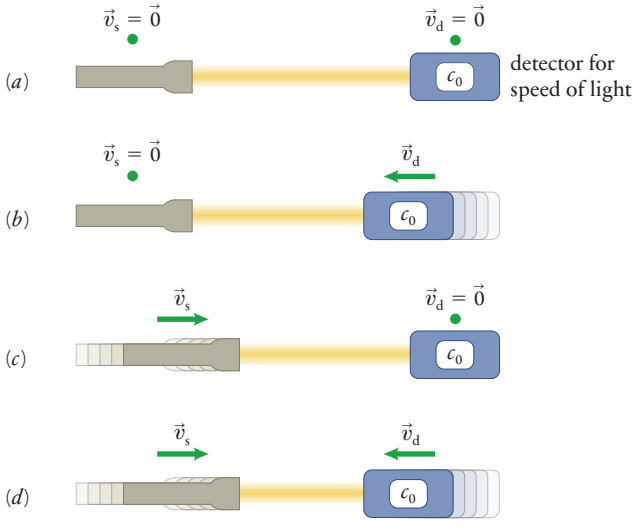
Example 14.3 reaffirms what we learned in Chapter 6: The speed of an everyday object relative to an observer depends on the motion of both the object and the observer. Relative to an observer moving along with the tube in Figure 14.9, the ball moves not at speed v_b but at speed $v_b + v_t$. Measurements of the **speed of light in vacuum,** denoted by c_0, however, show something very different:*

> **The measured value of the speed of light propagating in vacuum is c_0, independent of the motion of the source relative to the observer making the measurement.**

This counterintuitive result was first suspected after an experiment done by A.A. Michelson and E.W. Morley in 1887. Many experiments since then have provided extensive confirmation of the Michelson-Morley results. To help you grasp the full meaning of this remarkable result, imagine that we shine a beam of light from a source into a device that is at rest relative to the source and that measures the speed of light in vacuum (**Figure 14.10a**; such an apparatus exists). The result is a value of 299,792,458 m/s (see Section 1.3), which we shall round off to 3.00×10^8 m/s. If we now move the measuring device toward the source at speed v_d and measure the speed of light, we obtain not $c_0 + v_d$ but again c_0 (Figure 14.10*b*)! This baffling result is like discovering that the speed of the ball with respect to the tube in Figure 14.8 is the same regardless of the motion of the tube relative to Earth.

Experiments show that it makes no difference whether we move the source, move the measuring device, or move both. In all cases, the device measures the speed of light in vacuum to be c_0 (Figures 14.10*c* and *d*). This result holds even when the source or detector moves at a substantial fraction of c_0: If a source moves toward you at a speed of $0.2c_0$, you still measure the

Figure 14.10 The measured value of the speed of light in vacuum does not depend on the speeds of the source and the measuring device.



*The subscript zero on c_0 is there to remind us that this is the value obtained in vacuum. The zero indicates the absence of any medium (that is, substance) through which the light propagates.

CONCEPTS

light traveling toward you at c_0, not $c_0 + 0.2c_0 = 1.2c_0$. After more than 100 years of repeated experiments, we can most certainly rule out an error in measurement.

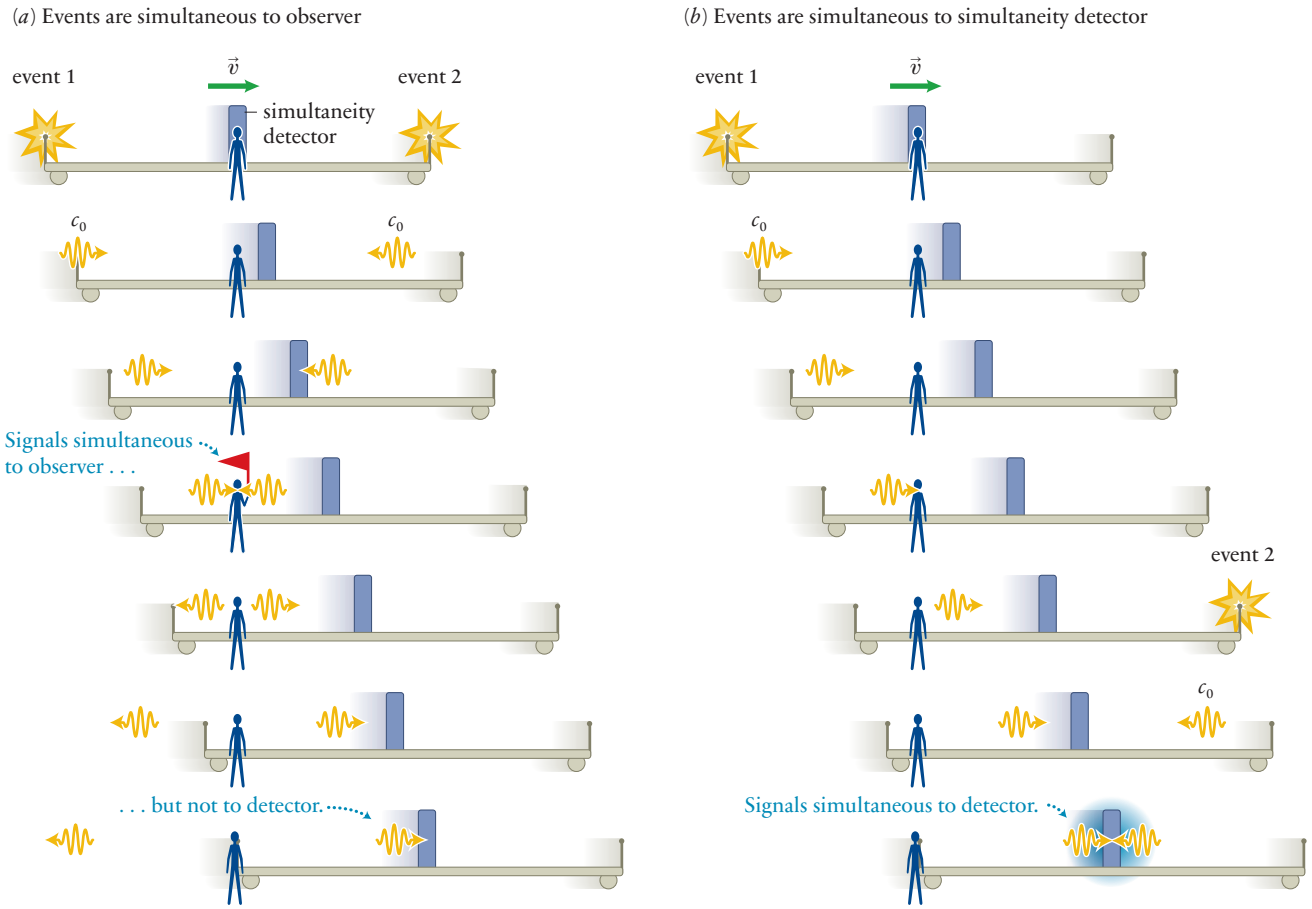
Any quantity whose value does not depend on the choice of reference frame is called an **invariant.** Because the speed of light is the same in all inertial reference frames, it is an invariant. Invariants play an important role in physics because we want to develop theories and laws that are the same in different reference frames. To construct these theories and laws, we first need to establish which physical quantities are invariants. The invariance of these quantities tells us something fundamental about the underlying laws of the universe.

Measuring the speed of light involves nothing more complicated than measuring the distance traveled by light in a certain time interval and dividing that distance by the time interval. The unsettling conclusion from Figure 14.10 is that measured values of distances and time intervals must somehow depend on the motions of the source and the reference frame. This conclusion forces us to reexamine what we mean by *distances* and *time intervals*, two of the most basic quantities we introduced back in Chapter 2.

Let us start with the experimental fact that the speed of light in vacuum is always exactly c_0 and examine the consequences this result has on the experiment of Figure 14.7*a*. Suppose the entire setup—the two light sources and the simultaneity detector midway between them—now moves at a very high speed v past an observer at rest on the ground, as in **Figure 14.11**. (By *very high* I mean a speed large enough for the detector to move an appreciable distance in the time interval it takes the light signals to reach the detector.) Suppose further that the observer sees the two emission events occur at the same instant; that is, he is midway between the two sources at the instant they emit a light signal, and so the two signals reach him at the same instant (middle panel in Figure 14.11*a*). In other words, *the events are simultaneous in the observer's reference frame.*

As you can see from Figure 14.11*a*, however, the simultaneity detector moves to the right as the signals travel toward it, and so it receives signal 2 before it receives signal 1. Therefore the blue light does not go on. In the reference frame of the detector, the two signals were *not* emitted simultaneously! Meanwhile, the observer is confused: In his reference frame, the two

Figure 14.11 If the sources and the simultaneity detector move together at high speed relative to an observer, then (*a*) events that are simultaneous to the observer will not be so to the detector. (*b*) To make the signals strike the simultaneity detector at the same instant, the source behind the detector must emit its signal before the source in front of the detector.



signals *were* emitted simultaneously, and so why didn't the blue detector light come on?

14.4 Suppose the observer in Figure 14.11*a*, at rest relative to the ground and standing midway between the two sources when they emit signals, is joined by another observer who moves along with the simultaneity detector. (*a*) Based on what the detector tells them, do the observers agree that the light signals do not reach the detector at the same instant? (*b*) How does each observer explain why the signals don't reach the detector at the same instant?

Suppose the sources in Figure 14.11 emit the signals in such a way that the fast-moving detector receives them at the same instant and the blue light turns on. To the observer at rest on the ground, source 1 must have emitted its light signal before source 2 emitted its signal because signal 1 has to catch up with the detector moving away from it (Figure 14.11*b*). In his reference frame, the two events are *not* simultaneous, and so he is once again confused by what the detector is "saying": Blue light on means the events were simultaneous. The blue light tells us that the detector sees exactly the situation depicted in Figure 14.7*a*, and so in the reference frame of the detector the emission events are simultaneous. From these experiments we infer that:

Events that are simultaneous in one reference frame need not be simultaneous in another reference frame that is moving relative to the first.

What does this relativity of simultaneity imply for the grid of synchronized clocks of Figure 14.6? To answer this question, consider a pair of synchronized clocks at rest relative to observer A midway between the two clocks in **Figure 14.12**. The clocks are mounted on opposite ends of a rod that is one grid unit long. Because the clocks are synchronized, they strike 3 o'clock at the same instant according to observer A. Now consider the situation from the point of view of observer B, who is moving to the left at constant speed relative to A.

Figure 14.12 Two synchronized clocks are unsynchronized in the reference frame of an observer moving relative to the clocks.

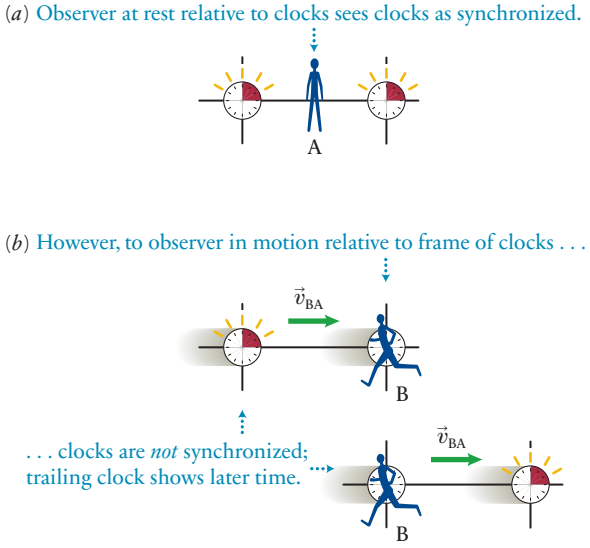
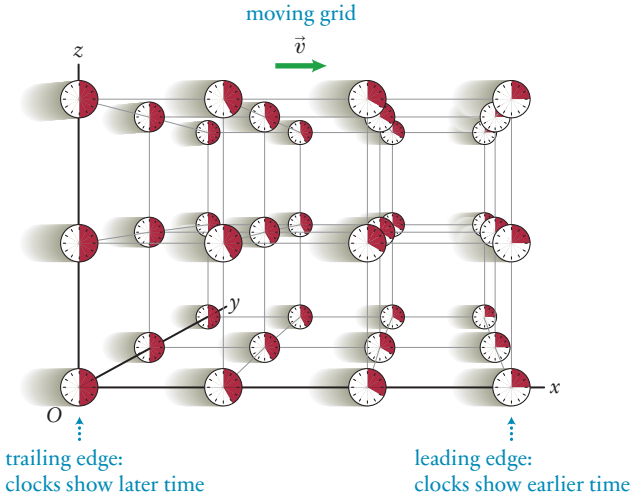


Figure 14.13 Desynchronization of time in a moving reference frame.



From B's perspective, the two-clock unit moves to the right, just like the simultaneity detector in Figure 14.11. Suppose events 1 and 2 in Figure 14.11 are our two clocks striking 3 o'clock. As Figure 14.11*b* shows, these two events occur at the same instant in the reference frame of the detector. Relative to observer B, however, the left clock strikes 3 o'clock (event 1) before the right clock does (event 2). So, in Figure 14.12*b*, too, the left clock strikes 3 o'clock *before* the right clock in the reference frame of observer B. In other words, the clock on the trailing end of the grid unit runs ahead of the clock on the leading end. This statement must hold true for any pair of neighboring clocks along the line of motion, which means that the clocks on a grid moving relative to an observer desynchronize as shown in **Figure 14.13**. If you were to move along with the grid, all clocks would strike, say, 3 o'clock at the same instant. If the grid moves relative to you, however, the clocks at the trailing end of the grid strike 3 o'clock before those at the leading end.

What is one instant in one reference frame is a multitude of different instants in a reference frame that is moving relative to the first reference frame. As we saw in analyzing Figures 14.11 and 14.12, events that are simultaneous in one reference frame are not simultaneous in a reference frame that is moving relative to the first one.* Simultaneity is a relative concept, and time intervals are therefore not invariants.

*The only events we consider here are ones that occur at different positions along the line of motion of two reference frames in motion relative to each other. Events occurring along an axis perpendicular to the line of motion can be simultaneous in both reference frames.

CONCEPTS

The reason this notion of the relative nature of simultaneity defies common sense is that these effects are observable only at speeds that are much greater than everyday speeds. In Figure 14.11*a*, for example, in the time interval it takes the light to travel half the distance between the two sources, the simultaneity detector travels one-fourth of that distance. In other words, the detector moves at half the speed of light, or 1.5×10^8 m/s, close to a million times faster than the speed of an airplane! Had it instead moved at the speed of an airplane, the displacement between the top and bottom pictures in Figure 14.11 would have been negligible, and so, for all practical purposes, the signals would have arrived simultaneously at the detector both according to an observer on the ground and according to one on the moving assembly. It is only when observers (or detectors) move at speeds close to c_0 relative to each other that deviations from our everyday experience cannot be ignored.

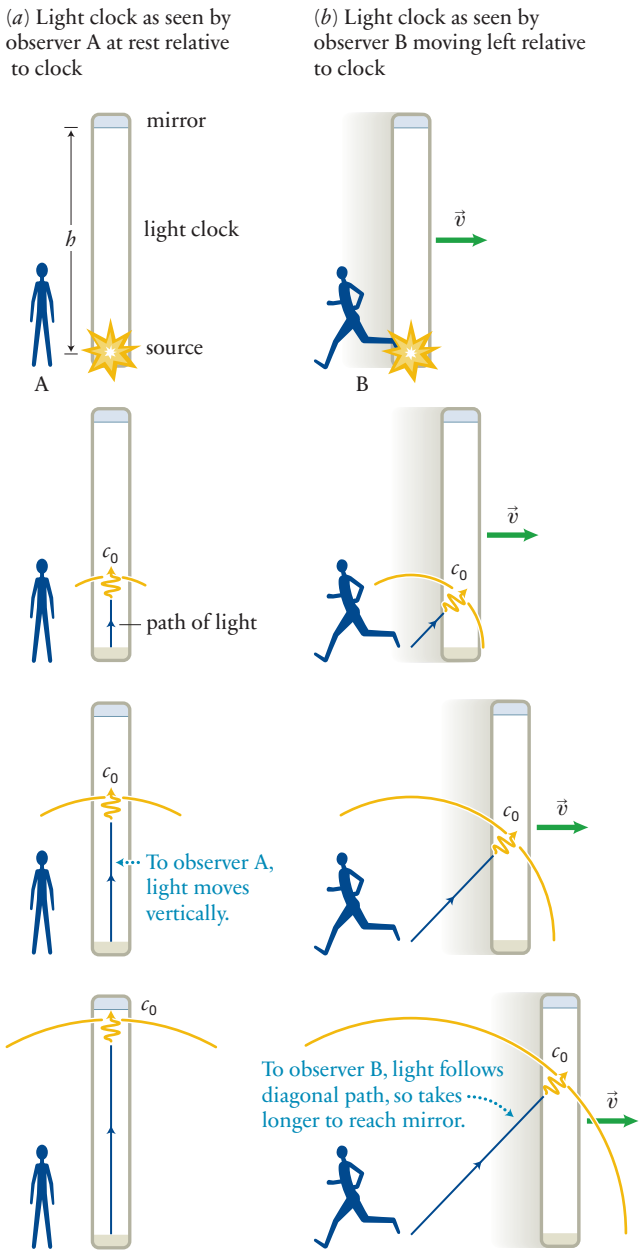
14.5 Suppose the setup in Figure 14.11 is 10 m long. (*a*) How long does it take a light signal to travel from source to detector? (*b*) If the setup moves at 20 m/s, how far does it travel in the time interval you calculated in part *a*? (*c*) By how much does the displacement of the detector shorten (or lengthen) the travel time of the light signals? (*d*) Do you expect the time interval you calculated in part *c* to be measurable? (*e*) Repeat parts *b*–*d* for a detector moving at 1.0×10^8 m/s.

14.3 Space-time

The fact that simultaneity is relative rather than absolute makes us rethink the whole concept of time. If simultaneity depends on the motion of observers relative to each other, shouldn't the rate at which a clock is observed to tick depend on the motion of the observer relative to the clock? The answer is yes, and here is the easiest way to see that. Consider the device shown in **Figure 14.14**, called a *light clock*. At the bottom is a source that emits a very short light signal; at the top is a mirror that reflects the signal back down. At the instant the reflected signal reaches the bottom, the source is triggered to emit a new light signal. The up-and-down motion of the signals determines the period of this clock. If the mirror is a distance h away from the source, it takes each light signal a time interval h/c_0 to travel from the source up to the mirror and the same time interval to travel back down to the source. The period of this clock is thus $2h/c_0$. Just as the ticking of a mechanical clock defines the unit of time for the clock, this period defines the unit of time for the light clock.

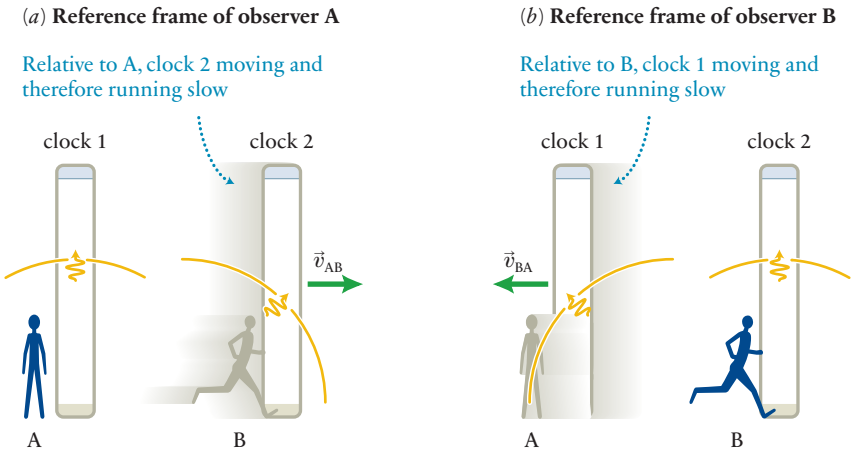
Consider this clock from the perspective of two observers: one at rest relative to the clock (observer A, Figure 14.14*a*) and one moving to the left at very high speed v relative to the clock (observer B, Figure 14.14*b*). To observer A the clock signals move straight up and down. To observer B, however, the signals move along a longer diagonal path. Because the signals travel at exactly c_0 according to both observers, observer B sees the signals take longer to travel from source to mirror and from mirror to source.

Figure 14.14 Because the speed of light in vacuum is always c_0, the light signal in a light clock takes longer to reach the mirror according to an observer moving relative to the clock than to an observer at rest relative to the clock.



CONCEPTS

Figure 14.15 Each observer sees the other observer's clock run more slowly.



14.6 (*a*) In Figure 14.14, is the period of the clock according to observer B greater than, equal to, or less than the period according to A? (*b*) Does the period of the clock according to B increase, decrease, or stay the same if we increase v?

Because the distance the signal travels from source to mirror and back to source is greater according to B than according to A, it takes longer for the clock to complete a cycle according to observer B. Therefore, the clock period is longer for B than it is for A. Observer B sees the clock as running more slowly than observer A. The greater the speed at which the clock moves relative to B, the more slowly she observes it ticking off its units of time. This effect is called **time dilation** because the units of time appear stretched out (dilated):

When a clock is in motion relative to an observer, that observer sees the time units of the clock as being longer than the time units of an identical clock at rest relative to her. Consequently, she measures the clock in motion as running slow relative to the clock at rest.

14.7 On your desk in front of you are two watches, one mechanical and one electronic. For each minute on the electronic watch, the mechanical watch advances by only 59 s. (*a*) Is the mechanical watch fast or slow relative to the electronic one? (*b*) Using the mechanical watch as the standard, is the electronic one fast or slow?

Ordinarily, if your watch runs slow relative to some clock, that clock runs fast relative to your watch, as you just saw in Checkpoint 14.7. Things are different when dealing with time dilation. Consider, for example, two identical clocks: clock 1, which is at rest relative to observer A, and clock 2, which is traveling along with observer B. To observer A, clock 2 is in motion and therefore running slow compared to clock 1 (**Figure 14.15*a***). To observer B, however, clock 2 is at rest and clock 1 is moving (Figure 14.15*b*). From B's perspective, it is now the signal in clock 1 that takes longer to reach its mirror. So to B, clock 1 appears to run slowly relative to clock 2. The effect of time dilation is thus symmetrical, as required by the principle of relativity: Each observer sees the other observer's clock run more slowly.

Note that in Figure 14.14*b* the light signal leaves the source when the clock is at one position and returns to the source when the clock is at a different position. This means that the time interval between these events measured in B's reference frame is not the proper time interval. Observer A measures the proper time interval because the emission and return take place at the same position relative to A. This means that the proper time interval between the two events is the shortest, and we can say:

An observer who measures the proper time interval between two events (that is, an observer for whom the events occur at the same location) obtains the shortest time interval between those two events.

EXAMPLE 14.4 My nap was shorter

You are flying in a plane, and a friend is on the ground. Who measures a shorter duration of (*a*) a nap you take and (*b*) a nap your friend takes?

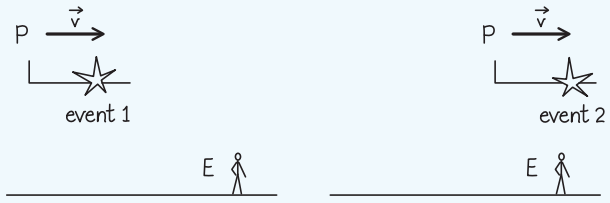
❶ **GETTING STARTED** Like Example 14.1, this problem involves two observers (my friend and I) and two pairs of events (the beginning and end of each person's nap). My friend is in the Earth reference frame (E); I am in the reference frame of the plane (P). To simplify the discussion, I denote the beginning and end of my nap as events 1 and 2, respectively, and the beginning and end of my friend's nap as events 3 and 4, respectively.

CONCEPTS

2 DEVISE PLAN The proper time interval between two events is the shortest time interval between them. Therefore this problem boils down to determining who measures the proper time interval for each nap. I should therefore make a sketch for each pair of events.

3 EXECUTE PLAN (*a*) **Figure 14.16** shows the beginning (event 1) and end (event 2) of my nap on the plane. To me, the two events take place at the same location, which means that the time interval I measure is the proper time interval. To my friend, the two events take place at different locations (I fall asleep far to his left and wake up directly overhead), so the interval he measures is not the proper one. Thus I measure the shorter duration for my nap. ✔

Figure 14.16



(*b*) **Figure 14.17** shows the beginning (event 3) and end (event 4) of my friend's nap on the ground. To my friend, the two events take place at the same location, which means that the time interval he measures is the proper time interval. To me, the two events take place at different locations (my friend falls asleep in front and below the plane and he wakes up directly underneath), so the interval I measure is not the proper one. Thus my friend measures the shorter duration for his nap. ✔

Figure 14.17



4 EVALUATE RESULT Each nap happens in the reference frame of the napper, and so it makes sense that each person measures the proper time interval of her or his nap.

You may ask yourself if this slowing down of moving clocks is just a peculiarity of the type of clock we have constructed. Who measures time with light clocks anyway? Suppose you have two identical mechanical clocks adjusted so that their periods are the same as those of a pair of identical light clocks. Now suppose you pair each mechanical clock with one of the light clocks and set one pair in motion. The stationary pair remains synchronized. Likewise, the moving pair must remain synchronized for this reason: If the motion of the clocks caused them to become unsynchronized, an observer moving along with them would note the difference in ticking rates and from this difference be able to determine his state of motion, which would violate the principle of relativity (see Section 6.3). Indeed, no one has ever observed a pair of initially synchronized clocks becoming unsynchronized because the pair's state of motion changes. So, we know that (1) moving light clocks run more slowly and (2) mechanical clocks behave just as light clocks do. We therefore conclude that moving mechanical clocks—or any other clock, for that matter, including your biological clock—must exhibit the same time-dilation effects as a light clock. It is time itself that dilates!

Even though time dilation is extremely small at everyday speeds, it has been measured by comparing two atomic clocks (which are not light clocks), one at rest on the surface of Earth and the other aboard an aircraft (see the box "Desynchronized twins" on page 472).

Direct evidence that the time interval between two events depends on the motion of the observer is provided by cosmic-ray muons, fast-moving elementary particles created when highly energetic protons from deep space collide with atoms some 10 km up in Earth's atmosphere. Laboratory experiments show that muons at rest have a *half-life* of 1.5×10^{-6} s. That is, if you have 100 muons at some instant $t = 0$, half of them will decay into some other type of particle during one half-life. Thus you will have only 50 muons left at $t = 1.5\ \mu$s, 25 left at $t = 3.0\ \mu$s, and so on.

Traveling at a speed nearly equal to c_0, a muon travels 450 m in 1.5×10^{-6} s, and so the number of muons should decrease by a factor of 2 each 450 m they travel through the atmosphere. Experiments done at the top of Mt. Washington at an altitude of 1900 m showed that 568 muons strike a certain detector per hour. At sea level the muons have traveled an additional 1900 m, or about four times 450 m, and so the count should be diminished by a factor of $\left(\frac{1}{2}\right)^4$, yielding approximately 35 muons per hour. When the same detector is placed at sea level, however, not 35 but 412 muons per hour are detected—more than ten times as many as expected. How can this be?

Because of the time dilation that occurs at the speed they travel, the muons' half-life of 1.5×10^{-6} s is stretched to 1.4×10^{-5} s in our reference frame at rest on the ground. So, in the Earth reference frame the muons exist long enough for a substantial fraction of them to reach the ground before disintegrating.

14.8 Could you lengthen your life by traveling at high speed?

Motion is a relationship between time and distance. Thus, now that we understand how relative motion

CONCEPTS

Desynchronized twins

Suppose one member of a set of twins embarks on a space trip, flying away from Earth at a speed of $0.98c_0$. Ten years after her departure, she turns around and returns to Earth, again traveling at speed $0.98c_0$. To observers on Earth monitoring the trip, clocks on the space ship run five times more slowly than clocks on Earth. Likewise, the twin's heart—a clock—beats five times more slowly. Consequently, by the time she returns to Earth, she has aged not 20 years but 4 years (20 years/5)! When she meets her twin brother again, who has remained Earthbound this whole time interval, he's 16 years older than she is. According to him, she's been gone 20 years, but to her, the trip lasted only 4 years.

This scenario is disturbing because it defies our commonsense notion of time. The experiment has been carried out with a pair of identical ("twin") atomic clocks, one flown Eastward twice around the world while the other sat at the airport. Upon its return, the flown clock was behind the airport clock, and the time difference matched what theory predicts.

This desynchronization of twins (or twin clocks) is sometimes called the *twin paradox* because viewing the situation from the perspective of the traveling twin appears to yield a contradiction: To her it is her brother on Earth who is moving at $0.98c_0$, and so *he* should be the one aging more slowly! The paradox is that, as explained earlier, she does not see him as being 16 years younger—both agree that she is the younger one once the trip is over. In other words, even though she sees Earth/brother moving at $0.98c_0$ and so should see Earth clocks running *slow,* these Earth clocks measure a *longer* duration for her trip than her own clocks. How can this be?

The key to resolving this paradox is to remember that the brother remains in an inertial reference frame during the entire trip, but the sister doesn't. The spaceship has to turn around to return to Earth, and this change in direction means acceleration. During this acceleration, a marble initially at rest on the floor of the spaceship does not remain at rest, and thus there is a time interval during the trip when she is in a noninertial reference frame. In contrast, a marble initially at rest on Earth's surface is not affected by the spaceship's acceleration: in other words, the Earth remains an inertial reference frame all the while the traveling twin is away. Thus the physical situation is different in the two reference frames: the Earth reference frame always inertial, the reference frame of the spaceship not always inertial. The theory of special relativity shows that during the time interval in which she is accelerating, the sister would see Earth clocks speed *up,* so that when she gets back to Earth, she is not at all surprised to see that her brother has aged much more than she has.

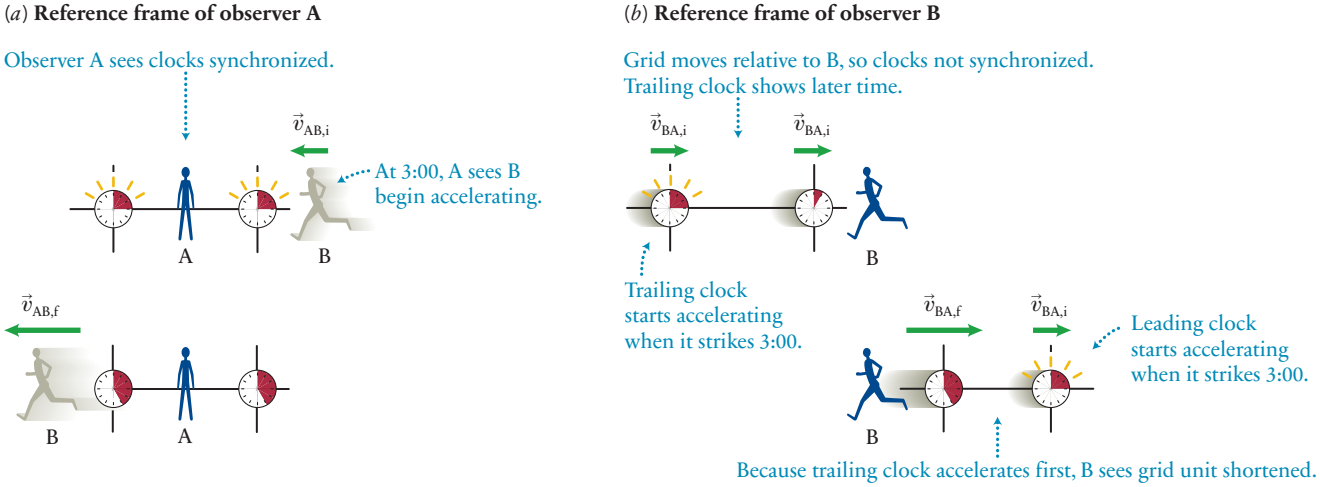
changes our ideas about time measurements, a logical next question to investigate might be whether motion changes the measurement of *distance*. In other words, is there a length change analogous to time dilation? An intriguing thought experiment shows us that the answer is yes. Picture yourself traveling along with some muons as they move from the summit of Mt. Washington to sea level. They are at rest relative to you, and so there is no time dilation in your reference frame: According to your clock, the muons have a (proper) half-life of 1.5×10^{-6} s. If you are traveling with the muons (at rest in their reference frame), you see the ground rushing toward you at a speed of nearly c_0. Thus it should still take 4 half-lives for the surface of the sea to travel the 1900 m toward you, and as a result you should again see about 35 muons by the time sea level reaches you. You see 412, though, just as before. Because you "know" that the muon half-life is 1.5×10^{-6} s, you are forced to conclude that the height of Mt. Washington has somehow shrunk!

Another way to see that length measurements are affected by motion is to imagine measuring the length of a *moving* car. One way would be to record the positions of the front and back of the car on the pavement and then measure the distance between those marks. You need to make sure that you record the two positions simultaneously, however, because the car is moving relative to you. If you record the back position at one instant and the front position some time later, you won't get a correct result. Remember, however, that simultaneity is a relative concept, and therefore we can expect length measurements also to depend on the motion of an observer relative to the object whose length the observer is measuring.

To see how length measurements are affected by motion, consider again a grid unit with a pair of clocks. The unit is at rest and the clocks are synchronized in the reference frame of observer A (**Figure 14.18*a***). Suppose a second observer B speeds up leftward when the clocks read 3 o'clock in A's reference frame. Because the clocks are synchronized, both of them read 3 o'clock in A's reference frame when observer B begins speeding up. From observer B's perspective, observer A and the clocks speed up to the right (Figure 14.18*b*). Because observer B is moving relative to the grid unit, the clocks are desynchronized, with the clock on the

CONCEPTS

Figure 14.18 Length contraction of an accelerating grid.



trailing end being ahead of the clock on the leading end (see Figure 14.13). In observer B's reference frame, therefore, the clock on the trailing end strikes 3 o'clock and begins speeding up *before* the clock on the leading end does. Consequently B sees the distance between the clocks (and the length of the rod between them) decrease. Whenever the unit speeds up in B's reference frame, B sees the length of the rod decrease. The faster the unit moves, the shorter the rod appears to B.

When the unit slows down relative to observer B, the opposite happens: The clock on the trailing end begins slowing down *before* the clock on the leading end begins slowing down, and so B sees the rod become longer again. When the unit is at rest relative to B, she measures the rod to have its **proper length:**

> **The length of an object measured by an observer at rest relative to the object is called the object's proper length.**

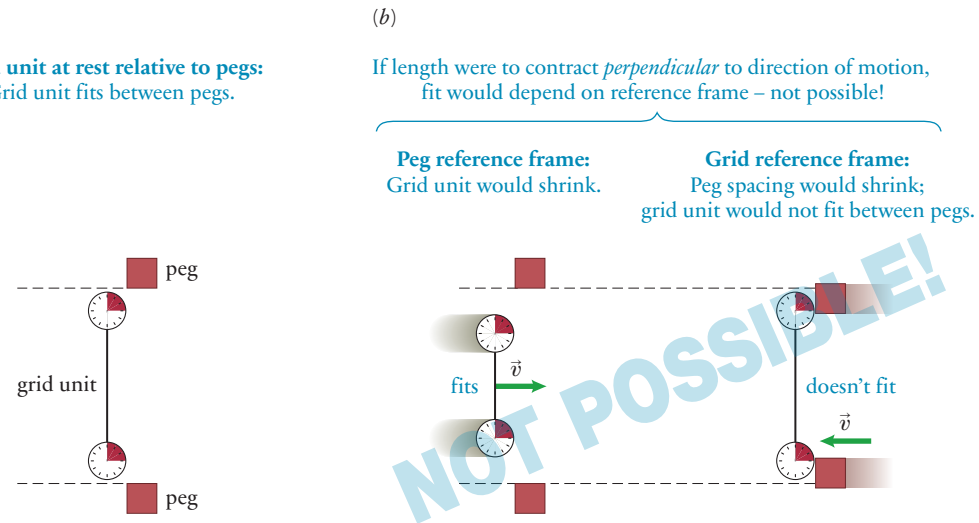
An observer measuring the length of an object moving relative to the observer always measures a value that is shorter than the object's proper length.

> **14.9** Suppose observer B in Figure 14.18 carries a unit of two synchronized clocks that is identical to the unit that is at rest relative to A. Observer B begins speeding up when the clocks in her reference frame read 3 o'clock. What happens to the length of B's unit according to observer A?

As this checkpoint shows, the shortening of the rod's length, like time dilation, is symmetrical, in agreement with the principle of relativity. Each observer sees the other observer's grid move at the same relative speed and therefore shortened by the same amount.

Note that the shortening occurs only in the direction of motion. To see why, consider our clock grid unit aligned as in **Figure 14.19**. With the unit at rest, pegs are

Figure 14.19 Length contraction cannot occur in a direction perpendicular to the direction of motion.



CONCEPTS

Common sense and special relativity

That observers who are in motion relative to each other assign different values to the length of an object and to the time interval between two events clashes with common sense. Even though experiments confirm the predictions of special relativity, it is hard to believe that, say, the length of a rod and the time interval between your lunch and dinner do not have fixed, absolute values. This difficulty arises from the fact that in everyday circumstances we can get away with treating lengths and time intervals as invariants.

How "real" then is length contraction or the slowing down of clocks? It is just as real as the pitch of an approaching car horn sounding higher than that of a stationary or receding car. (This phenomenon is called the *Doppler effect;* see Section 17.7.) Riding along in a car with a stuck car horn, you measure the same pitch a roadside observer would measure when the car is at rest. Yet there is no denying that the pitch is perceived to change when car and observer are in motion relative to each other. These observations don't mean that the horn changes when it moves relative to an observer. What changes is the pitch *perceived by the observer*.

The same statements can be made about a moving rod. Any observer measuring the rod's length when the rod is at rest relative to him obtains the same value as you would (the proper length). Just as the pitch of a moving car horn is perceived to be changing, so is the measured length of a moving rod. The rod itself doesn't change—the *measurement of its length* yields a different value. In that sense relativistic effects are every bit as real as the changing pitch of a horn. For example, few muons would make it across what we measure to be 10 km of atmosphere if that length did not appear to be shortened to less than 1 km to them.

placed above and below it so that it fits snugly between them (Figure 14.19*a*). If the unit were to get shorter when it is in motion, it would fit more easily between the pegs (Figure 14.19*b*). Now imagine observing the same situation while traveling along with the unit. The unit is now at rest from your point of view and so does not shorten. The pegs, however, are in motion relative to you, and so the distance between them becomes smaller and the clocks hit the pegs. However, whether or not the clocks hit the pegs cannot depend on the reference frame of the observer, and so there cannot be any shortening in a direction perpendicular to the direction of motion.

When we discussed time dilation and considered the moving light clock in Figure 14.14*b*, we tacitly assumed its height to be fixed. Now we see that this assumption was justified.

The shortening of objects in the direction of motion is called **length contraction:**

> **When an object moves relative to an observer, the observer measures the length of the object in the direction of its motion to be shorter than the proper length of the object.**

Length contraction is a direct consequence of the relativity of simultaneity. Returning to the cosmic-ray muons, we can now see why they reach Earth's surface. To them, Earth's atmosphere flies by at nearly the speed of light. Length contraction reduces the thickness of the atmosphere so much that the atmosphere passes by the muons before they decay. To us on the ground, time dilation lengthens the half-life of the muons, and so they can cross the proper length of the atmosphere. To them, length contraction decreases the distance to the ground, and so they reach the ground during their proper lifetime.

Time and space are connected. If we consider time and space to constitute a four-dimensional *space-time* entity, we see that, depending on the reference frame, the time dimension stretches (dilates) and one of the three spatial dimensions contracts. As we shall see in Section 14.5, a quantity called the *space-time interval* remains unaffected by this distortion of space-time geometry. At everyday speeds, these distortions are so small that they are not perceived. For motions at very high speeds, such as that of cosmic-ray muons, however, the consequences of these effects can be significant (see the box "Common sense and special relativity").

CONCEPTS

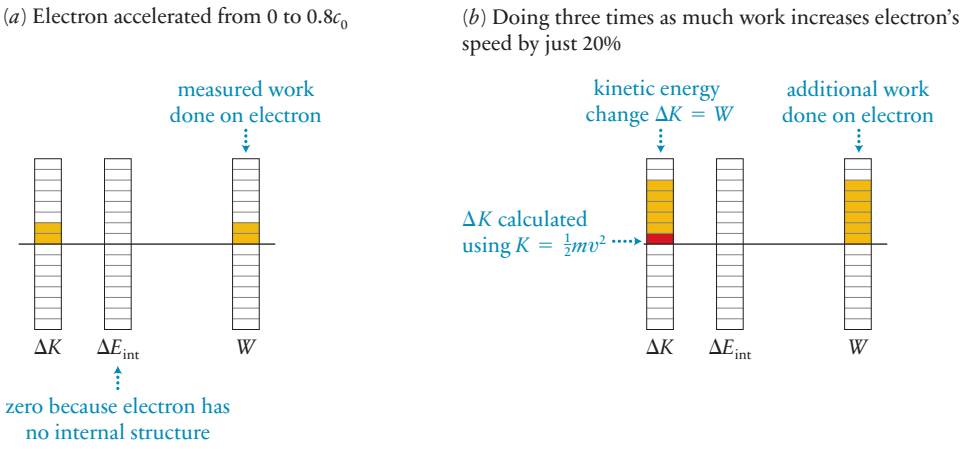
14.10 Consider a long pole moving through a tunnel, with the direction of motion being along the long axis of the pole. An observer on the ground determines that the moving pole has the same length as the tunnel. Does an observer moving along with the pole agree that the pole fits exactly in the tunnel?

14.4 Matter and energy

We developed the entire framework of mechanics from our everyday notions of space and time. In Chapters 2 and 3, we introduced the definitions of displacement, velocity, and acceleration. Then we introduced inertia, which depends on a ratio of changes in velocity (see Section 4.5). Many other quantities, such as momentum, energy, and force, in turn depend on inertia. So all the physics we have developed so far hinges on our concepts of space and time.

As we just saw in the preceding sections, however, everyday notions of space and time fail at very high

Figure 14.20 Simplified energy diagrams for accelerating an electron.



speed. Additionally, there is plenty of experimental evidence that our notion of energy fails at very high speeds. For example, experiments show that quadrupling the kinetic energy of an electron initially moving at $0.8c_0$ by doing work on it increases its speed by just 20% rather than doubling it as we would expect from the expression $K = \frac{1}{2}mv^2$.

Figure 14.20*a* shows the energy diagram for accelerating an electron from rest to $0.8c_0$. (The diagram is simplified by lumping all internal energy into one bar.) Because electrons have no known internal structure, $\Delta E_{int} = 0$, and so only their kinetic energy can change by doing work on them. The change in kinetic energy must thus be equal to the work done on the electron.

Now suppose we do additional work on the electron after it has attained a speed of $0.8c_0$. If the additional work is three times the amount of work done in accelerating the electron from rest to $0.8c_0$, our expression for kinetic energy, $K = \frac{1}{2}mv^2$, predicts a new speed that is twice the original speed: $v_f = 1.6c_0$. Experiments show, however, that the speed increases by just 20% and the value we obtain for ΔK using the expression $K = \frac{1}{2}mv^2$ is very much smaller than expected (Figure 14.20*b*). We must therefore conclude that our expression for kinetic energy is not valid at high speeds.

This experiment shows that it is more difficult to accelerate an electron when it is moving at $0.8c_0$ than when it is at rest. Now remember that an object's inertia is a measure of its resistance to acceleration (Section 4.2). We have so far assumed that an object's inertia is constant—an intrinsic property that is independent of the object's physical state or motion relative to an observer. Experimental data show that this assumption does not hold as an object's speed approaches c_0. In particular, as its kinetic energy increases, the object's inertia increases.



14.11 Do two observers, one moving at very high speed relative to the other, agree on the inertia of an object? Why or why not?

As Checkpoint 14.11 shows, inertia depends on the reference frame and therefore is not an invariant. In Chapter 13 we introduced mass—the property that determines the gravitational interaction of objects. The way in which objects interact gravitationally cannot depend on the motion of an observer relative to the interacting objects, which means that, in contrast to inertia, mass must be an invariant. That is to say, all observers should agree on the value of the mass of an object, regardless of their motion relative to the object. That mass is an invariant and inertia is not implies that there *is* a difference between mass and inertia, contrary to what we established in Chapter 13. As we shall see in Section 14.7, the difference between the two becomes apparent only at very high speeds.

Let us now turn to the nonkinetic part of energy: internal energy. Remember that the internal energy of an object or system is a quantitative measure of the state of that object or system. Because the state of an object cannot depend on the motion of the observer relative to the object, internal energy also must be an invariant.

So, we see that kinetic energy and inertia are not invariants, but internal energy and mass are. We also found that inertia depends on kinetic energy. Experiments show a similar connection between mass and internal energy. The primary example is provided by nuclear reactions, such as those taking place in the Sun. One of these reactions is the fusion of a proton and a deuteron D (the nucleus of a hydrogen isotope consisting of one neutron and one proton) to form a helium-3 nucleus (two protons and one neutron), while 8.80×10^{-13} J is released as radiation energy:

$$p + D \rightarrow {}^3He + \text{radiation}.$$

CONCEPTS

Figure 14.21 Simplified energy diagram for the fusion of a proton and a deuteron. During the reaction, nuclear energy is emitted in the form of radiation (identified as R). Experiments show that the amount of energy released is 8.80×10^{-13} J.

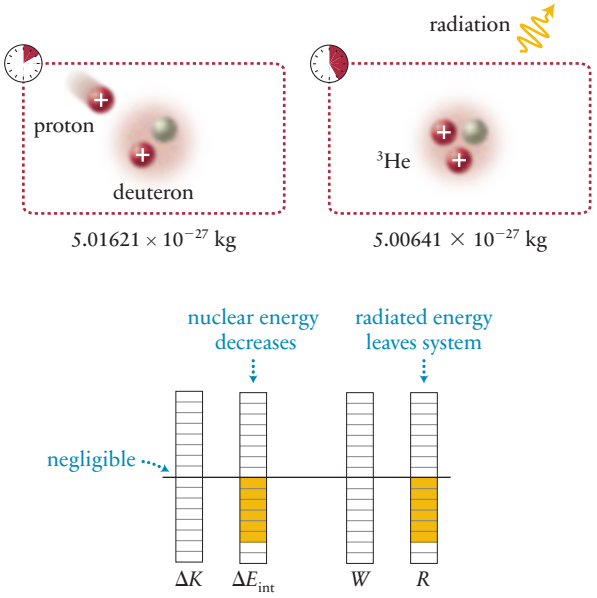


Figure 14.22 An ionized deuterium molecule and a helium ion contain the same elementary particles, yet have different masses. (Nuclei not to scale.)

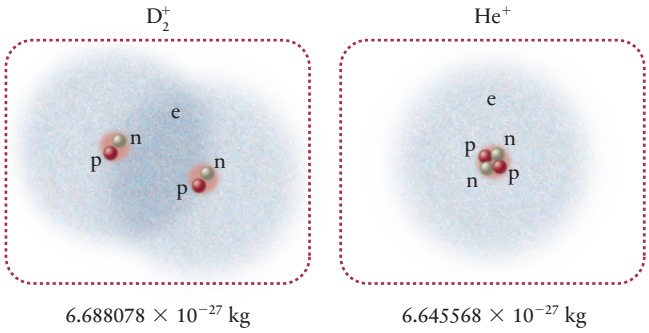


Figure 14.21 shows a simplified energy diagram for this reaction. Given that this is a nuclear reaction, nuclear energy (a form of internal energy) must be converted to radiation energy. Because the radiation energy leaves the system, the internal energy decreases.

Measurements show that the sum of the mass of the proton and the mass of the deuteron ($m_p + m_D = 5.01621 \times 10^{-27}$ kg) is greater than the mass of the helium-3 nucleus (5.00641×10^{-27} kg). In other words, the mass of the system decreases as its internal energy decreases.

Another example of the connection between mass and internal energy is provided by comparing the masses of an ionized deuterium molecule and an ionized helium atom (**Figure 14.22**). Both consist of two protons, two neutrons, and an electron, but the two particles have different configurations. Accurate measurements show that this different in configuration—which means a difference in internal energy—results in a small mass difference.

These observations indicate a relationship between mass and internal energy that is parallel to the one between inertia and kinetic energy:

Increasing the internal energy of a system increases the mass of the system.

Mass is a property we associate with particles or objects that have physical substance—made of what we call "matter." Internal energy is a quantitative measure of the object's state. Now we see that the two are proportional to each other and that whenever a system's state changes, both its internal energy and its mass change.

14.12 In each of the following situations, does the mass of the system increase, decrease, or stay the same? (*a*) A spring is compressed (system: spring). (*b*) A cup of coffee cools down (system: coffee).

EXERCISE 14.5 The mass of energy

(*a*) In the nuclear reaction illustrated in Figure 14.21, what amount of mass is "lost"? (*b*) What is the ratio of energy decrease to mass loss as a result of the reaction? (*c*) Show that the units of your answer to part *b* are equivalent to m^2/s^2.

SOLUTION (*a*) The mass decreases by $(5.00641 \times 10^{-27}\ \text{kg}) - (5.01621 \times 10^{-27}\ \text{kg}) = -9.80 \times 10^{-30}$ kg. ✔

(*b*) The text states that the energy released in the reaction is 8.80×10^{-13} J. Therefore the ratio of energy decrease to mass loss is $(8.80 \times 10^{-13}\ \text{J})/(9.80 \times 10^{-30}\ \text{kg}) = 8.98 \times 10^{16}$ J/kg. ✔

(*c*) $\text{J/kg} = (\text{kg} \cdot \text{m}^2/\text{s}^2)/\text{kg} = \text{m}^2/\text{s}^2$. ✔

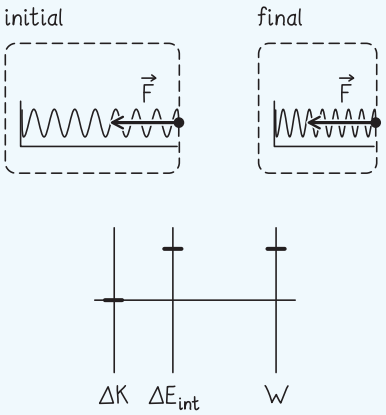
EXAMPLE 14.6 Spring mass

A spring of spring constant $k = 0.16$ N/m is compressed to half its 50-mm relaxed length. Using the result of Exercise 14.5, determine the change in the spring's mass.

❶ **GETTING STARTED** I begin by drawing a simplified energy diagram for the compressed spring (**Figure 14.23**). The work done on the spring during the compression increases the spring's potential energy and therefore the spring's internal energy. I need to determine the increase in mass corresponding to this energy increase.

CONCEPTS

Figure 14.23



2 DEVISE PLAN To calculate the increase in the spring's potential energy, I can use Eq. 9.23, $\Delta U_{\text{spring}} = \frac{1}{2}k(x - x_0)^2$. In part *b* of Exercise 14.5, I determined the ratio of energy change to mass change. I can use this ratio to convert the spring's energy change to a mass change.

3 EXECUTE PLAN The spring has been compressed by an amount $x - x_0 = 50\text{ mm} - 25\text{ mm} = 2.5 \times 10^{-2}\text{ m}$. From Eq. 9.23, I get $\Delta U_{\text{spring}} = \frac{1}{2}(0.16\text{ N/m})(2.5 \times 10^{-2}\text{ m})^2 = 5.0 \times 10^{-5}\text{ J}$. In Exercise 14.5*b*, I found that the ratio of energy change to mass change is $8.98 \times 10^{16}\text{ J/kg}$, so the change in the spring's mass is $(5.0 \times 10^{-5}\text{ J})/(8.98 \times 10^{16}\text{ J/kg}) = 5.6 \times 10^{-22}\text{ kg}$. ✔

4 EVALUATE RESULT The change in the spring's mass is much too small to be measured. That is reassuring because in my experience I've never been able to observe any change in mass when compressing springs.

Exercise 14.5 and Example 14.6 show that the ratio of mass change to energy change is very small; that is, a tiny amount of mass corresponds to an enormous amount of energy. Table 14.1 lists the fractional changes in mass for a number of changes in state (corresponding to changes in energy). As a rule of thumb, nuclear reactions release an amount of energy corresponding to about 10^{-3} of the mass; chemical reactions release an amount of energy corresponding to about 10^{-9} of the mass. As you can see from the table, other processes cause even smaller fractional changes in mass. For example, no one has been able to observe the change in mass that occurs when we heat a cup of tea because this change is on the order of 10^{-13} kg, much too small to measure.

In Chapter 13 we established that mass and inertia are equivalent at everyday speeds. Consequently increasing the internal energy of an object at rest must increase not only the object's mass but also its inertia. So, whereas mass depends on internal energy only, inertia depends on both internal energy and kinetic energy:

> **Increasing the energy $E = K + E_{\text{int}}$ of a system increases its inertia. The increase in inertia is the same whether the added energy goes into kinetic energy K, internal energy E_{int}, or a combination of the two.**

Table 14.2 summarizes the correspondence between matter and energy. The mass and internal energy of a system are invariants—that is, independent of the motion of observer and system relative to each other. Inertia and energy are not invariants. In addition, energy is a conserved quantity. In the second part of this chapter, we'll use the conservation principles to derive quantitative expressions for the relationships between mass and internal energy and between inertia and energy.

Table 14.1 Changes in energy and fractional mass changes due to changes in state

System before	System after	Energy change (J)	Fractional change in system's mass
1-kg book and Earth	1-kg book and Earth separated by an additional 1 m	10	2×10^{-41}
cup of tea	cup of tea heated by 10 °C	4000	5×10^{-13}
1 mole of H_2 molecules	2 moles of H atoms	5×10^{5}	3×10^{-9}
p + D	^{3}He	9×10^{-13}	2×10^{-3}

Table 14.2 Matter and energy

Quantity	Invariant?	Conserved?	Related to
E_{int}	yes	no	mass
E	no	yes	inertia

14.13 Do the inertia and mass of the system increase, decrease, or remain the same when (*a*) an elementary particle is accelerated (system: particle) and (*b*) an elementary particle slams into a metal target (system: particle and target)?

CONCEPTS

Self-quiz

1. Imagine you are at the origin of Figure 14.6, where, as in Exercise 14.2, a light pulse takes 1 s to travel one grid unit. When your clock reads 12:00:05, what do you observe as the reading on the clocks at positions (3, 0) and (0, 2)?
2. Which of the following statements are true when we compare a moving light clock and one at rest in the Earth reference frame? (*i*) The signal in the moving clock travels a greater distance during one cycle. (*ii*) The signal in the moving clock takes more time to complete one cycle. (*iii*) The cycle on the moving clock is longer. (*iv*) Time on the moving clock runs more slowly. (*v*) The moving clock completes an hour first.
3. Standing near the finish line at a race track, you measure the time interval between two cars as they cross the finish line. Is the time interval between these events as measured by the driver in the second car the same as, greater than, or less than the time interval you measure?
4. A particle of mass m is moving at speed v toward an identical particle that is at rest in the Earth reference frame. Is the inertia of the particles in their zero-momentum reference frame greater than, equal to, or less than m?
5. An observer on the ground observes a car moving past a gate and notes that the car is longer than the gate. For an observer in the car, which of the following are possibilities? (*i*) The gate is longer than the car. (*ii*) The gate is the same length as the car. (*iii*) The gate is shorter than the car.

Answers

1. Because it takes 1 s for light to travel one grid unit, it takes 3 s for the clock at (3, 0) to be seen by you. Its time is thus 3 s earlier than that of the origin clock, or 12:00:02. The light from the clock at (0, 2) takes 2 s to reach you, and so that clock reads 12:00:03.
2. All but (*v*) are true. Because the signal in the moving clock travels an equal vertical distance but a greater horizontal distance, the overall distance is greater. At constant speed, a greater distance requires a longer time interval. This longer time interval makes the duration of the clock's cycle longer, and the clock runs slow. If the clock runs slow, the hour passes less quickly.
3. Because you measure both events at the same position, you obtain the proper time interval between them. The driver in the second car moves relative to you, and so, according to the argument used in the solution to Example 14.4, measures a longer time interval.
4. In the zero-momentum reference frame both particles are moving, and therefore each particle has an inertia that is greater than the mass m.
5. Because moving objects contract in the direction of motion, the gate gets shorter to the observer in the car. Because the gate is shorter than the car in the Earth reference frame, it is also shorter in the car's reference frame. Thus only (*iii*) is possible.

14.5 Time dilation

Let us determine how the period of a light clock depends on the motion of the clock relative to the observer. The time interval between the instant a light signal is emitted at the bottom of a light clock and the instant the signal returns to the bottom is the period of that clock. In the reference frame of observer A in **Figure 14.24*a***, the clock of height h is at rest, and so the period measured by observer A is $T_A = 2h/c_0$. In the reference frame of observer B, the clock is in motion and the signal follows the inverted-V path shown in Figure 14.24*b*. Because this path is longer than $2h$, the period T_B measured by observer B is longer than T_A.

During one round trip of its light signal, the clock moves a distance vT_B to the right in B's reference frame. From the right triangle in Figure 14.24*b*, we see that the distance d traveled by the signal during that time interval must satisfy

$$\left(\frac{d}{2}\right)^2 = h^2 + \left(v\frac{T_B}{2}\right)^2. \tag{14.1}$$

Because the signal travels at speed c_0, we have $d = c_0T_B$. Substituting this relationship into Eq. 14.1 and multiplying the result by 4 to remove any fractions, we get

$$(c_0T_B)^2 = 4h^2 + (vT_B)^2 \tag{14.2}$$

or

$$(c_0^2 - v^2)T_B^2 = 4h^2. \tag{14.3}$$

Dividing both sides of this equation by c_0^2 gives

$$\frac{c_0^2 - v^2}{c_0^2}T_B^2 = \frac{4h^2}{c_0^2} = T_A^2, \tag{14.4}$$

where we have used $T_A = 2h/c_0$. Rearranging terms, we get

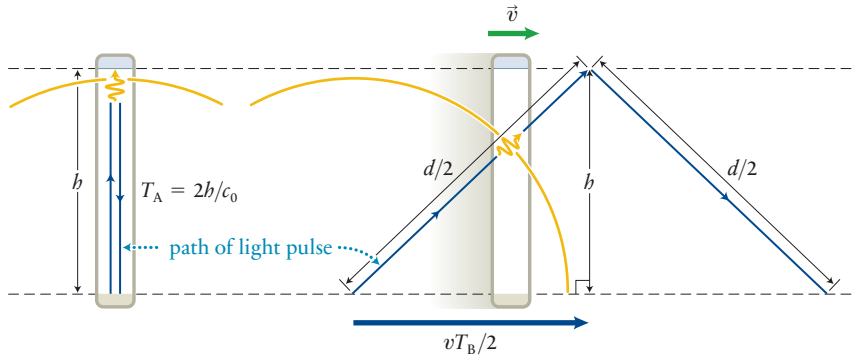
$$T_B = \sqrt{\frac{c_0^2}{c_0^2 - v^2}}T_A = \frac{1}{\sqrt{1 - v^2/c_0^2}}T_A. \tag{14.5}$$

Because the factor in front of T_A is greater than 1 for $0 < v < c_0$, we see that $T_B > T_A$, as expected.

Figure 14.24 The period of a light clock moving relative to an observer is determined by the time interval it takes the light signal to travel the two diagonal distances $d/2$.

(*a*) **Reference frame of observer A:** clock at rest relative to observer

(*b*) **Reference frame of observer B:** clock moving relative to observer



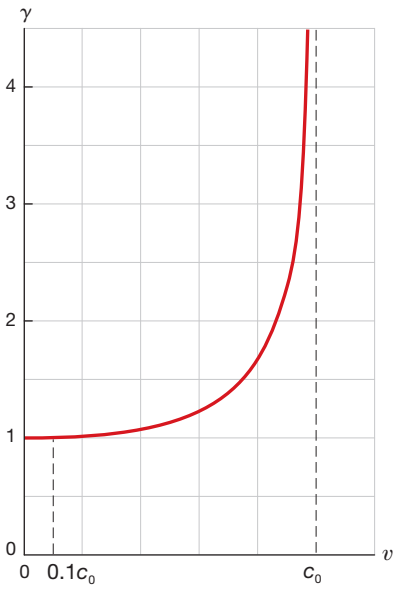


Figure 14.25 As the speed v at which any object travels approaches c_0, the Lorentz factor γ approaches infinity. Up to a speed of about $0.1c_0$, γ is very close to 1.

Table 14.3 Gamma factor

v/c_0	γ
0	1
0.1	1.005
0.5	1.15
0.9	2.3
0.98	5.0
0.99	7.1
0.999	22
0.9999	71
0.99999	220
0.999999	710

The factor on the right side of Eq. 14.5 is often represented by the Greek letter gamma, γ, and called the **Lorentz factor:**

$$\gamma \equiv \frac{1}{\sqrt{1 - v^2/c_0^2}}. \tag{14.6}$$

This factor is a measure of how pronounced relativistic deviations are from our everyday notions of time and space. At ordinary speeds, $v \ll c_0$, and so $\gamma \approx 1$; at high speeds, γ approaches infinity, as shown in **Figure 14.25**.

Table 14.3 lists values of γ for some values of v/c_0. Note that γ remains very close to 1 even up to very high speeds. For example, an object traveling at $0.1c_0$ could circle Earth in about 1 s, and yet the deviation of γ from 1 is less than 1%, meaning that the effects of time dilation are very small. As a rule of thumb, we can say that for speeds less than $0.1c_0$ the deviation of γ from 1 is negligible. We'll refer to speeds less than $0.1c_0$ as *nonrelativistic speeds* and speeds greater than $0.1c_0$ as *relativistic speeds*.

Conversely, we can obtain the ratio v^2/c_0^2 from γ by first squaring Eq. 14.6:

$$\gamma^2 = \frac{1}{1 - v^2/c_0^2} \tag{14.7}$$

and then inverting this equation:

$$\frac{1}{\gamma^2} = 1 - \frac{v^2}{c_0^2}, \tag{14.8}$$

so

$$\frac{v^2}{c_0^2} = 1 - \frac{1}{\gamma^2}. \tag{14.9}$$

For nonrelativistic speeds, we can approximate γ as*

$$\gamma = \frac{1}{\sqrt{1 - v^2/c_0^2}} \approx 1 + \tfrac{1}{2}\frac{v^2}{c_0^2} \quad (v/c_0 \ll 1), \tag{14.10}$$

so

$$\gamma - 1 \approx \tfrac{1}{2}\frac{v^2}{c_0^2} \quad \text{(within two significant digits when } v < 0.1c_0\text{)}. \tag{14.11}$$

As you will see shortly, this equation is useful for calculating γ when $v < 0.1c_0$.

14.14 Taking each motion to be relative to the Earth reference frame, calculate γ for (*a*) a car moving at 30 m/s, (*b*) an airplane moving at 250 m/s, (*c*) a spacecraft moving at 10,000 m/s, and (*d*) an electron moving at $0.60c_0$.

Using the factor γ, we can simplify Eq. 14.5 to

$$T_B = \gamma T_A. \tag{14.12}$$

*From the binomial expansion $[\sqrt{1+\epsilon}\,]^{-1} = 1 - \frac{1}{2}\epsilon + \frac{3}{8}\epsilon^2 + \cdots$ (see Appendix B). For small ϵ, we can ignore terms of order ϵ^2 and higher.

QUANTITATIVE TOOLS

This equation provides an exact quantitative expression for time dilation. In words, the period T_B measured by an observer in motion relative to the clock is lengthened by a factor γ relative to the period T_A measured by an observer at rest relative to the clock. Note that T_B is not a proper time interval between the emission of successive light signals because these events occur at different positions in the reference frame of the observer. Only an observer at rest relative to the clock can measure the proper time interval between the emission of successive light signals. To this observer, the clock is at rest, and so the light signal travels a distance $2h$ between successive emission events. The proper time interval between the emission of successive signals in the clock is thus $2h/c_0$.

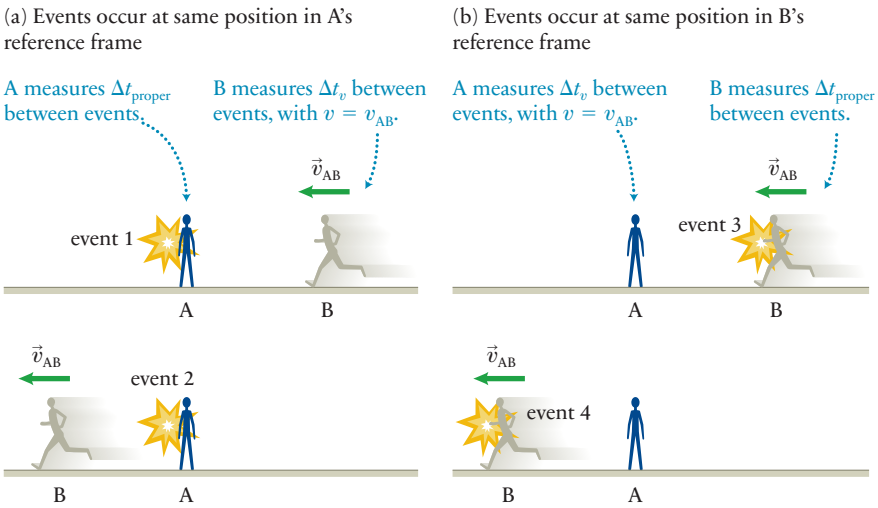
Time dilation does not occur in light clocks only. As we pointed out in Section 14.3, the principle of relativity tells us that any clock should exhibit the same effects. Time dilation affects the measurement of the time interval between any two events and therefore time itself. Generalizing Eq. 14.12, we can say that the time intervals between two events as measured by two observers in motion relative to each other are related by

$$\Delta t_v = \gamma \Delta t_{\text{proper}}, \tag{14.13}$$

where Δt_{proper} is the time interval measured by an observer for whom both events occur at the same position, Δt_v is the time interval between the same two events measured by a second observer moving at speed v relative to the first observer, and γ is the factor given by Eq. 14.6.

Note that Eq. 14.13 involves two observers measuring the time interval Δt between two events. When applying this equation, you must therefore always begin by identifying the two events and the two observers. Next you should determine which observer measures the proper time interval and what speed v to use in Eq. 14.13. As an example, consider **Figure 14.26**. In Figure 14.26*a* two events occur at the same position relative to observer A, and so A measures the proper time interval Δt_{proper}. Observer B moves at speed v_{AB} relative to A, and so she measures a time interval $\Delta t_v = \gamma \Delta t_{\text{proper}}$, where the v in γ is equal to v_{AB}. In Figure 14.26*b* two events occur at the same position relative to B, and so B measures the proper time interval Δt_{proper}. Observer A moves at speed $v_{BA} = v_{AB}$ relative to B and so he measures a time interval $\Delta t_v = \gamma \Delta t_{\text{proper}}$, where the v in γ is equal to v_{AB}.

Figure 14.26 Proper time measurements.



EXAMPLE 14.7 Flight delay

You are on an airplane flying at a steady 250 m/s at cruising altitude. At noon the pilot announces that you are passing over the center of London and exactly 6.0 h later, according to your watch, comes on the address system to say that you are passing over Abu-Dhabi. (*a*) Do observers on the ground measure a longer, equal, or shorter time interval between these two events? (*b*) Is the difference between the time interval measured on the ground and that measured by your watch detectable?

❶ **GETTING STARTED** This problem involves two reference frames (the Earth reference frame and the reference frame of the plane) and two events (plane passing over London and plane passing over Abu-Dhabi). I denote the plane passing over London as event 1 and the plane passing over Abu-Dhabi as event 2. The time interval between these events is 6.0 h in the reference frame of the plane.

❷ **DEVISE PLAN** The proper time interval is the shortest time interval between the events, and so this problem boils down to determining who measures the proper time interval. Once I have determined that, I can use Eq. 14.13 to determine the time interval measured by the ground observers and determine the difference between the two intervals.

❸ **EXECUTE PLAN** (*a*) Relative to the plane, events 1 and 2 occur at the same position. Therefore the 6.0-h time interval I measure is the proper time interval in the plane reference frame. The ground observers move at 250 m/s relative to my reference frame, and so, according to Eq. 14.13, they measure a longer time interval. ✔

(*b*) To determine the difference, I subtract the proper time interval from both sides of Eq. 14.13:

$$\Delta t_v - \Delta t_{\text{proper}} = \gamma \Delta t_{\text{proper}} - \Delta t_{\text{proper}} = (\gamma - 1)\Delta t_{\text{proper}}.$$

From Eq. 14.11 I have

$$\gamma - 1 \approx \tfrac{1}{2}\frac{v^2}{c_0^2} = \frac{(250\ \text{m/s})^2}{2(3.00 \times 10^8\ \text{m/s})^2} = 3.47 \times 10^{-13}.$$

The difference in the time intervals is thus $(3.47 \times 10^{-13})(6.0\ \text{h})(3600\ \text{s}/1\ \text{h}) = 7.5 \times 10^{-9}$ s, which is orders of magnitude smaller than can be detected with ordinary clocks. ✔

❹ **EVALUATE RESULT** That the difference in time intervals is too small to be easily measured is reassuring in two ways. First, from experience I know that my watch doesn't stop telling the correct time while flying (other than time zone adjustments). Second, the speed of the airplane is just a small fraction of the speed of light $(250\ \text{m/s})/(3.00 \times 10^8\ \text{m/s}) \approx 10^{-6}$, and so, for all practical purposes, $\gamma = 1$.

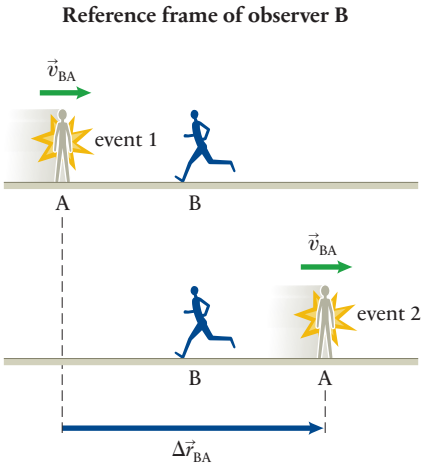


Figure 14.27 Observer B can determine the proper time interval between any pair of events by measuring the time interval between them and the distance separating them.

Let's examine Figure 14.26*a* from observer B's point of view and see if there is any way that B can determine the proper time interval between the two events without the need to refer to observer A's reference frame. In B's reference frame, the events occur at two locations (**Figure 14.27**). Let's choose the x axis such that we can denote A's displacement in B's reference frame by Δx. Relative to B, A moves at velocity $\vec{v}_{BA}$. If B measures a time interval Δt between the two events, then the x component of A's displacement is

$$\Delta x = v_{BAx}\Delta t = v_{BAx}\gamma \Delta t_{\text{proper}}, \tag{14.14}$$

where we have used Eq. 14.13 to convert Δt to the proper time interval between the events. Equation 14.14 relates the proper time interval to measurements taken entirely in B's reference frame (Δx and Δt), but the equation also contains the velocity of A's reference frame relative to B. To obtain an expression that does not contain this velocity requires a bit of algebraic manipulation on the expression $(c_0\Delta t_{\text{proper}})^2$. We begin by multiplying and dividing this quantity by γ^2, substituting Eq. 14.8 for $1/\gamma^2$:

$$(c_0\Delta t_{\text{proper}})^2 = \gamma^2(c_0\Delta t_{\text{proper}})^2\left(1 - \frac{v^2}{c_0^2}\right). \tag{14.15}$$

Multiplying out the term in parentheses, we get

$$(c_0\Delta t_{\text{proper}})^2 = c_0^2\gamma^2(\Delta t_{\text{proper}})^2 - v^2\gamma^2(\Delta t_{\text{proper}})^2. \tag{14.16}$$

Because $\Delta t = \gamma \Delta t_{\text{proper}}$ (Eq. 14.13), the first term on the right is equal to $(c_0\Delta t)^2$. The magnitude of velocity $\vec{v}_{BA}$ is equal to the speed v that appears in the Lorentz factor, and so we can write $v_{BAx}^2 = v^2$. Equation 14.14

then tells us that the second term on the right in Eq. 14.16 is equal to $(\Delta x)^2$, and so we can rewrite this equation as

$$(c_0\Delta t_{\text{proper}})^2 = (c_0\Delta t)^2 - (\Delta x)^2. \tag{14.17}$$

Note that the right side of this expression contains only measurements made in B's reference frame: the time interval Δt and the spatial separation Δx of the events. The quantity $c_0\Delta t$ represents the distance traveled by a light signal in the time interval between the events. The quantity

$$s^2 \equiv (c_0\Delta t)^2 - (\Delta x)^2 \tag{14.18}$$

is called the **space-time interval** between two events. Because both terms on the right in Eq. 14.18 have units m^2, we see that s^2 represents a distance squared.

Depending on how $c_0\Delta t$ compares to Δx, we can distinguish between different types of event pairs. When the distance $c_0\Delta t$ traveled by a light signal in the time interval between the events is greater than the distance between the events, $(c_0\Delta t)^2 > (\Delta x)^2$, and Eq. 14.18 tells us that $s^2 > 0$. For such a pair of events, there is an inertial reference frame for which the two events occur at the same position ($\Delta x = 0$) and the events are separated in time only. For this reason the space-time interval of a pair of events for which $s^2 > 0$ is said to be *timelike*. A clock in the reference frame in which the two events occur at the same position measures the proper time interval Δt_{proper} between the events and uniquely determines their sequence in time. Because the speed of light c_0 and the proper time interval Δt_{proper} between two events have definite values, any observer who measures the distance and time interval between the two events must obtain the same space-time interval $s^2 = (c_0\Delta t_{\text{proper}})^2$ between them. In other words, the space-time interval is an invariant, even though neither Δx nor Δt is an invariant.

In a reference frame where the two events do not occur at the same position, $\Delta x > 0$. When $s^2 > 0$, Eq. 14.17 tells us that, for any reference frame in which $\Delta x > 0$, the time interval Δt between the events is greater than Δt_{proper}, as we concluded earlier. When Δt_{proper} is positive, Δt must be positive in all reference frames, and so a sequence of events that has a timelike space-time interval is the same in all inertial reference frames.

When the distance $c_0\Delta t$ traveled by a light signal in the time interval between the events is less than the distance Δx between them, Eq. 14.18 tells us that $s^2 < 0$. For such a pair of events, there is no inertial reference frame for which the two events occur at the same position and therefore the events have no proper time interval. There is, however, a reference frame for which the events are simultaneous ($\Delta t = 0$) and separated in space only: $(\Delta x)^2 = -s^2$. The space-time interval for a pair of events for which $s^2 < 0$ is therefore said to be *spacelike*. In any other reference frame, the events are not simultaneous, and we cannot uniquely determine their sequence. Depending on the reference frame, Δt can be positive, zero, or negative, and so either event may precede the other.

When two events are exactly separated by the distance that light travels in the time interval between them, $s^2 = 0$ and the space-time interval is said to be *lightlike*. When you look at a source of light (your desk lamp or a star, say), the emission of the light and your observation of that light are separated by a lightlike space-time interval.

Because the space-time interval s^2 is an invariant, it is timelike, spacelike, or lightlike regardless of the reference frame. How does the nature of the space-time interval affect the principle of causality we discussed in Chapter 1? To distinguish cause and effect, it is necessary to determine the

sequence of events. For a pair of events that have a timelike space-time interval ($s^2 > 0$), there is a reference frame in which we can study both events at the same position, and so we can determine their sequence with certainty and establish their causal relationship. For a pair of events that has a spacelike space-time interval ($s^2 < 0$), we cannot ascertain the time order of events and therefore cannot determine any causal relationship. For two events to be causally related, they must be able to interact physically. When $s^2 < 0$, the events occur in such rapid succession that a light signal cannot travel from one event to the other in the time interval between them. Because no information can travel faster than the speed of light, such events cannot interact and so cannot be causally related. Only events that have a timelike space-time interval can be causally related, and for such events the time sequence can be uniquely determined.



14.15 An elementary particle is observed to disintegrate after traveling 120 m in 5.0×10^{-7} s from the location where it was created. How long does the particle exist before disintegrating according to an observer at rest relative to the particle?

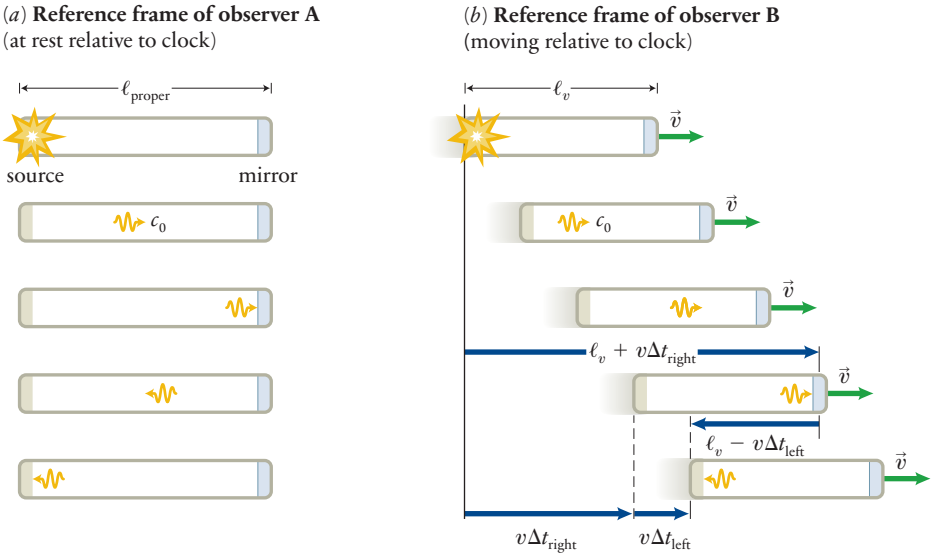
14.6 Length contraction

In Section 14.3 we saw that the length of an object measured by an observer moving relative to the object is shortened. Remember also that time dilation and length contraction are two manifestations of the same effect. To work out the amount of shortening, we now turn the clocks from Figure 14.24 on their side as in **Figure 14.28**, which shows the clock from the perspective of an observer A at rest relative to the clock and from the perspective of an observer B in motion relative to the clock.

We denote the proper length of the clocks by ℓ_{proper}. Our goal is to see what shortened length ℓ_v observer B obtains. According to A, the period of the clock is

$$T_{\text{A}} = \frac{2\ell_{\text{proper}}}{c_0}. \tag{14.19}$$

Figure 14.28 By calculating the distance traveled by the light pulse in the reference frame of the observer moving relative to the clock, we can work out by how much the length of the clock is shortened.



The period of the clock measured by B is given by Eq. 14.12: $T_B = \gamma T_A$. This period is the sum of two time intervals: the time interval Δt_{right} during which the light signal travels to the right from source to mirror and the time interval Δt_{left} during which the signal travels from mirror to source. During Δt_{right}, the clock moves a distance $v\Delta t_{right}$ as measured by B. The distance traveled by the signal during that time interval is thus $d_{right} = \ell_v + v\Delta t_{right}$. Because the signal travels at speed c_0 independent of the motion of the source or the motion of the observer, the distance traveled by the signal in the time interval Δt_{right} must be $d_{right} = c_0\Delta t_{right}$. Equating the two expressions for d_{right}, we have

$$c_0\Delta t_{right} = \ell_v + v\Delta t_{right}. \tag{14.20}$$

Solving for Δt_{right} yields

$$\Delta t_{right} = \frac{\ell_v}{c_0 - v}. \tag{14.21}$$

While the signal in the clock returns from the mirror to the source, the clock moves an additional distance $v\Delta t_{left}$.* The signal therefore travels a distance $d_{left} = \ell_v - v\Delta t_{left}$ again as measured by observer B. Because the signal travels at speed c_0 independent of the motion of the source or observer, the distance traveled by the signal in the time interval Δt_{left} must be $d_{left} = c_0\Delta t_{left}$. Equating the two expressions for d_{left}, we have

$$c_0\Delta t_{left} = \ell_v - v\Delta t_{left}. \tag{14.22}$$

Solving for Δt_{left} yields

$$\Delta t_{left} = \frac{\ell_v}{c_0 + v}. \tag{14.23}$$

Combining this result with Eq. 14.21, we get for the sum of the two time intervals

$$\begin{aligned} T_B &= \Delta t_{right} + \Delta t_{left} \\ &= \ell_v\left(\frac{1}{c_0 - v} + \frac{1}{c_0 + v}\right). \end{aligned} \tag{14.24}$$

Combining terms, we obtain

$$T_B = \frac{2\ell_v c_0}{c_0^2 - v^2}. \tag{14.25}$$

Dividing the numerator and denominator on the right by c_0^2, we get

$$T_B = \frac{2\ell_v}{c_0}\left(\frac{1}{1 - v^2/c_0^2}\right) = \frac{2\ell_v}{c_0}\gamma^2, \tag{14.26}$$

where we have substituted γ^2 from Eq. 14.7. Substituting this expression for T_B and Eq. 14.19 for T_A into Eq. 14.12, we get

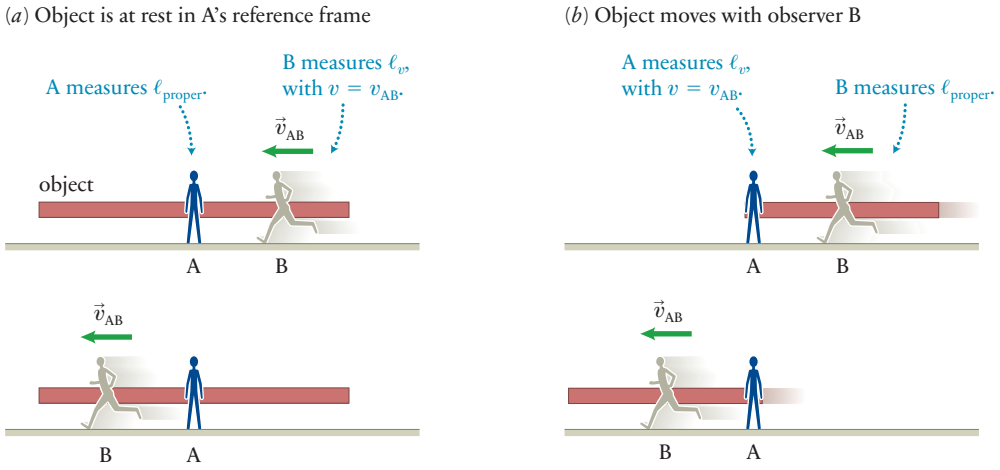
$$\frac{2\ell_v}{c_0}\gamma^2 = \gamma\frac{2\ell_{proper}}{c_0} \tag{14.27}$$

or

$$\ell_v = \frac{\ell_{proper}}{\gamma}. \tag{14.28}$$

*The subscript *left* refers to the motion of the light, not the motion of the clock.

Figure 14.29 Length measurements.



Because $\gamma > 1$, the measured length ℓ_v of the clock moving relative to us (or any other object moving relative to us) is shorter than its proper length ℓ_{proper}.

Note that Eq. 14.28 involves two observers measuring the length of an object: observer A at rest relative to the object and observer B in motion relative to it. When applying this equation, you must therefore always begin by identifying (1) the observers and (2) the observer's reference frame, if any, in which the object is at rest. The observer in this reference frame measures the object's proper length. As an example, consider the two observers in **Figure 14.29**. In Figure 14.29*a* the object is at rest in A's reference frame, and so A measures the proper length ℓ_{proper}. Observer B moves at speed v_{AB} relative to A, and so he measures a length $\ell_v = \ell_{proper}/\gamma$, where the v in γ is equal to v_{AB}. In Figure 14.29*b* the object is at rest relative to B, and so B measures the proper length ℓ_{proper}. Observer A moves at speed $v_{BA} = v_{AB}$ relative to B, and so she measures a length $\ell_v = \ell_{proper}/\gamma$, where the v in γ is equal to v_{AB}.

The Procedure box "Time and length measurements at very high speeds" summarizes how to treat time intervals and distances when dealing with motion at very high speeds.

PROCEDURE: Time and length measurements at very high speeds

At very high speeds, measurements of lengths and time intervals depend on the reference frame. Problems involving objects or observers moving at very high speeds therefore require extra care when we consider lengths and time intervals.

1. Identify each observer or reference frame with a letter and determine in which reference frame each numerical quantity given in the problem statement is measured.
2. If the problem involves any time intervals, identify the events that define that time interval. Determine in which observer's reference frame, if any, the two events occur at the same position. This observer measures the proper time interval between the events (Figure 14.26).
3. Identify any distances or lengths mentioned in the problem. Determine for which observer, if any, the object defining that distance or length remains at rest. This observer measures the proper length (Figure 14.29).
4. Use Eqs. 14.13 and 14.28 to determine the time intervals and lengths in reference frames other than the ones in which the intervals and lengths are given.

QUANTITATIVE TOOLS

EXAMPLE 14.8 Electron dash

An electron moving at $0.80c_0$ relative to the Earth reference frame travels the 100-m length of a building. What is the length of the building according to an observer moving along with the electron?

❶ **GETTING STARTED** This problem involves two reference frames: the reference frame of the building (which is the Earth reference frame) and the reference frame of the electron. Although the problem doesn't mention this explicitly, the 100-m length of the building is likely to be the length measured in Earth's reference frame, and so it is the building's proper length. I summarize this information in a table:

	Electron reference frame	Earth reference frame
v_{electron}	0	$0.8c_0$
ℓ_{building}	?	100 m

❷ **DEVISE PLAN** Equation 14.28 relates the building length measured by the observer in the electron's reference frame to the proper length. I know the building's proper length (100 m) and the speed of the observer relative to the reference frame in which the proper length is measured ($0.80c_0$). I can calculate γ from Eq. 14.6 and then use Eq. 14.28 to obtain the length of the building in the electron's reference frame.

❸ **EXECUTE PLAN** $\gamma = 1/\sqrt{1 - v^2/c_0^2} = 1/\sqrt{1 - (0.80)^2} = 1.67$, and so $\ell_v = \ell_{\text{building}} = (100\text{ m})/1.67 = 60\text{ m}$. ✔

❹ **EVALUATE RESULT** From our length contraction discussion in Section 14.3, I know that an observer measuring the length of an object that is moving relative to her measures a length shorter than the proper length. In this example, the observer is in the electron's reference frame, and the building being measured is moving relative to that reference frame. Therefore I expect the measured length to be shorter than the proper length.

EXAMPLE 14.9 Cosmic ray longevity

A proton strikes an air molecule in Earth's atmosphere, creating a muon 6.0 km above the ground. The muon travels toward Earth at nearly c_0 and decays just before reaching the ground. An observer traveling along with the muon determines that it lives for 2.2×10^{-6} s. According to this observer, what is the distance between the ground and the muon at the instant the muon is created?

❶ **GETTING STARTED** The problem mentions two reference frames: the Earth reference frame and the reference frame of the muon. The altitude is measured in the Earth reference frame, whereas the muon's lifetime is measured in the reference frame of the muon. The muon's lifetime is the time interval between its creation at an altitude of 6.0 km (event 1) and its decay at the ground (event 2). The speed of the muon in the Earth reference frame is (nearly) c_0. I summarize this information in tabular form:

	Earth reference frame	Muon reference frame
Δt		2.2×10^{-6} s (proper time interval)
h_{atm}	6.0 km (proper length)	?
v_{muon}	c_0	

❷ **DEVISE PLAN** The 6.0-km distance traveled by the muon in the Earth reference frame is the proper length of the atmosphere. To calculate the (contracted) distance traveled by the muon in the muon reference frame, I can use Eq. 14.28, but I first need to determine γ for the motion of the muon. I can determine the time interval between events 1 and 2 in the Earth reference frame by dividing the distance traveled in the Earth reference frame (6.0 km) by the muon speed, which I take to be c_0. This time interval is the lifetime of the muon in the Earth reference frame. Because I know the muon lifetime in the reference frame of the muon (2.2×10^{-6} s), which is the proper time interval, I can use Eq. 14.13 to determine γ.

❸ **EXECUTE PLAN** The distance traveled in the Earth reference frame is 6.0×10^3 m, and so in this reference frame the time interval between events 1 and 2 is $\Delta t_v = (6.0 \times 10^3\text{ m})/(3.00 \times 10^8\text{ m/s}) = 2.0 \times 10^{-5}$ s. In the muon reference frame, the muon is at rest and exists for the proper time interval $\Delta t_{\text{proper}} = 2.2 \times 10^{-6}$ s. I can obtain γ from Eq. 14.13:

$$\gamma = \frac{\Delta t_v}{\Delta t_{\text{proper}}} = \frac{2.0 \times 10^{-5}\text{ s}}{2.2 \times 10^{-6}\text{ s}} = 9.1.$$

Using this value for γ, I determine the distance ℓ_v traveled by the muon in its reference frame from Eq. 14.28:

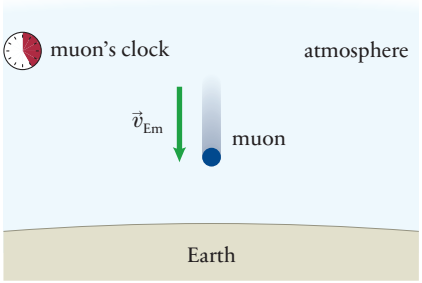
$$\ell_v = \frac{\ell_{\text{proper}}}{\gamma} = \frac{6.0 \times 10^3\text{ m}}{9.1} = 6.6 \times 10^2\text{ m}. \checkmark$$

❹ **EVALUATE RESULT** The muon lifetime in the Earth reference frame is longer than the muon's proper lifetime, as it should be. Because the muon lives longer in the Earth reference frame, it can travel a greater distance (6.0 km instead of 660 m). I can also check my result by considering the situation from the muon's reference frame. In that reference frame, the atmosphere flies by at nearly the speed of light for a time interval corresponding to the muon's proper lifetime, and so $\ell_v = c_0 \Delta t_{\text{proper}} = (3.00 \times 10^8\text{ m/s})(2.2 \times 10^{-6}\text{ s}) = 6.6 \times 10^2$ m, which is the result I got.

Figure 14.30 A muon created at the top of the atmosphere is able to reach the ground before decaying—but observers in the Earth reference frame and the muon reference frame would explain this result differently.

(*a*) **Earth reference frame**

Muon reaches ground because its clock runs slow.



(*b*) **Muon reference frame**

Muon reaches ground because atmosphere moves relative to it and hence is contracted.

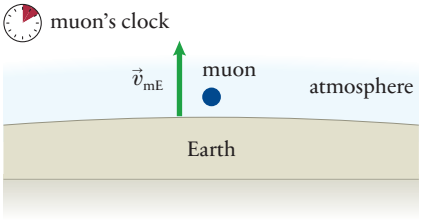


Table 14.4 Muon math

Reference frame	ℓ (km)	Δt (s)
Earth	6.0	2.0×10^{-5}
muon	0.66	2.2×10^{-6}

Figure 14.30 is a graphical representation of Example 14.9. To the observer in the Earth reference frame, the length of the muon's path is equal to the thickness (proper length) of the atmosphere, but the particle's lifetime is lengthened by time dilation (Figure 14.30*a*). Figure 14.30*b* shows the same decay viewed by the observer attached to the muon. To this observer, the muon is at rest and decays in the proper time interval, but the moving atmosphere is contracted. What appears as time dilation of the muon's lifetime to one observer appears as length contraction of the atmosphere to the other (**Table 14.4**). Time dilation and length contraction are thus two manifestations of the same effect.

In Chapter 6 we derived the Galilean transformation equations (Eqs. 6.4 and 6.5) to relate the coordinates of an event e in any two inertial reference frames A and B moving at constant velocity relative to each other. At very high speeds, these equations have to be replaced with equations that take into account the effects of time dilation and length contraction. For two reference frames A and B that have origins that coincide at $t_A = t_B = 0$ and that are moving at constant velocity in the x direction relative to each other, the new equations, called the **Lorentz transformation equations,** are (see Appendix E):

$$t_{Be} = \gamma\left(t_{Ae} - \frac{v_{ABx}}{c_0^2}x_{Ae}\right) \tag{14.29}$$

$$x_{Be} = \gamma(x_{Ae} - v_{ABx}t_{Ae}) \tag{14.30}$$

$$y_{Be} = y_{Ae} \tag{14.31}$$

$$z_{Be} = z_{Ae}. \tag{14.32}$$

At nonrelativistic speeds, these equations reduce to the Galilean transformation equations.

From the Lorentz transformation equations, we obtain that the x component of the velocity of an object o in reference frame A, v_{Aox}, is related to that in reference frame B, v_{Box}, by the expression (see Appendix E)

$$v_{Box} = \frac{v_{Aox} - v_{ABx}}{1 - \dfrac{v_{ABx}}{c_0^2}v_{Aox}}. \tag{14.33}$$

For nonrelativistic speeds, the denominator is approximately equal to 1 and we obtain the usual addition of velocities: $v_{Box} = v_{Aox} - v_{ABx}$ (see Eq. 6.9).

14.16 A car travels 1.0 km down a straight road at 100 km/h. How much shorter is the road in the car's reference frame?

14.7 Conservation of momentum

In Chapter 4 we introduced inertia to account for velocity changes in collisions and took the momentum of a particle to be the product of its inertia and its velocity. This approach led us to conservation of momentum, one of the fundamental principles that we have used throughout this text. In Section 14.4 we found that inertia depends on energy and therefore on speed. Because conservation principles are fundamental and because they embody an appealing simplicity, we would like to preserve them in light of the modifications required by special relativity. To this end, we begin again by taking the momentum of a particle as the product of its inertia and its velocity but this time allowing the inertia to depend on velocity:

$$\vec{p} \equiv m_v\vec{v}. \tag{14.34}$$

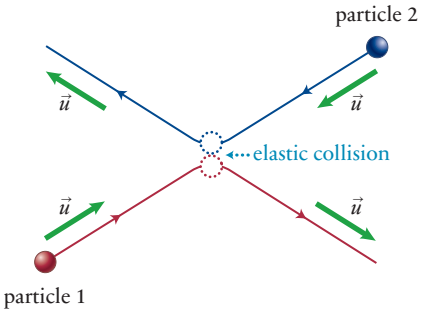
The subscript v on the inertia symbol m_v is there to remind us that inertia is no longer an invariant. We now examine what modifications are needed to our concept of inertia to preserve conservation of momentum.

To obtain an expression for m_v, let us look at an elastic collision between two identical particles and require the system to be isolated so that the momentum of the two particles remains constant. The particles are projected at each other at speed u (**Figure 14.31**). Their velocities are equal in magnitude but opposite in direction, both before and after the collision. The position of the center of mass of this system is at the location where the particles collide and it remains at rest. In other words, Figure 14.31 represents the collision viewed from the zero-momentum reference frame. We begin our analysis of this collision by choosing our axes as shown in **Figure 14.32*a***.

In the zero-momentum reference frame, the speeds of the particles are the same before and after the collision, and so we gain no information on how inertia depends on speed. Let us therefore consider the same collision from the point of view of an observer A moving parallel to the x axis of Figure 14.32 with a velocity equal to the component of the velocity of particle 1 along the x axis. To this observer, particle 1 moves up and down parallel to the y axis (Figure 14.32*b*). For convenience, let's denote the event when particle 1 leaves observer A as the *throw* and its return to observer A as the *catch*.

To an observer B moving along the x axis with a velocity equal to the component of the velocity of particle 2 along the x axis, the collision looks exactly the same except that particles 1 and 2 are interchanged (Figure 14.32*c*).

Figure 14.31 Collision between two identical particles as seen in the zero-momentum reference frame.



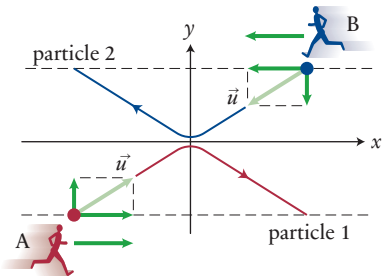
14.17 In Figure 14.32, is the time interval between the throw and the catch of particle 1 in reference frame A greater than, equal to, or less than the time interval between the throw and the catch of particle 2 in reference frame B?

Note that in all three reference frames only the y components of $\vec{v}_1$ and $\vec{v}_2$ change. Because the observers' motion is always in the x direction, there is no length contraction in the y direction, and so to both observers the particles travel the same distance d in the y direction between throw and catch.

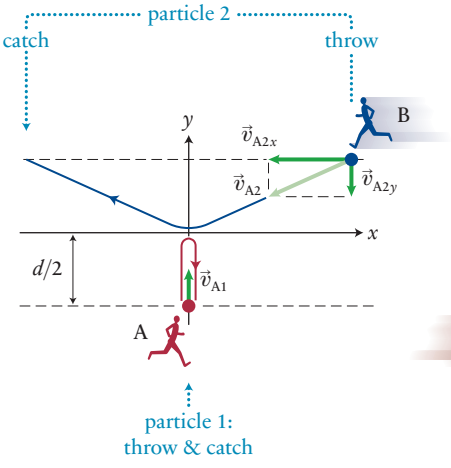
In reference frame A, the y component of the velocities of the two particles have the same magnitudes, but the x components are different. Therefore particles 1 and 2 move at different speeds v_1 and v_2, respectively, and so even though we began by declaring the two particles to be identical, the inertia m_{v1} of particle 1 measured in reference frame A must

Figure 14.32 Analysis of the trajectories in Figure 14.31 from three reference frames, showing our choice of axes.

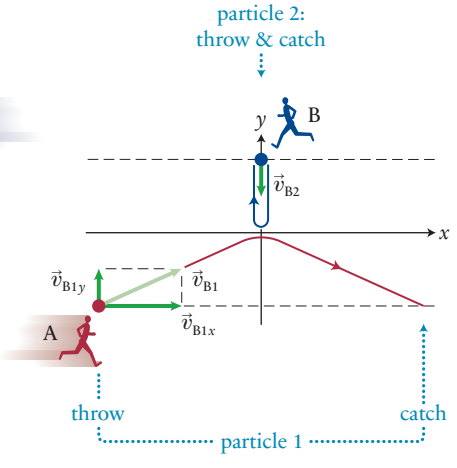
(*a*) **Zero-momentum reference frame**



(*b*) **Observer A's reference frame**



(*c*) **Observer B's reference frame**



be different from the inertia m_{v2} of particle 2 measured in this frame. We can obtain the ratio of these inertias by requiring the momentum of the two-particle system to remain constant. To this end, let us evaluate the velocity changes caused by the collision in reference frame A. Let the time intervals between the throw and catch of particles 1 and 2 in reference frame A be Δt_{A1} and Δt_{A2}, respectively. The changes in the y component of the velocities in reference frame A are thus

$$\Delta v_{A1y} = -\frac{d}{\Delta t_{A1}} \quad \text{and} \quad \Delta v_{A2y} = \frac{d}{\Delta t_{A2}}. \tag{14.35}$$

The time interval Δt_{A1} measured by observer A between the throw and catch of particle 1 in reference frame A is the proper time interval between these two events because they occur at the same position according to observer A: $\Delta t_{A1} = \Delta t_{\text{proper}}$. The time interval Δt_{A2} observer A measures between the throw and catch of particle 2, however, is not a proper time interval because for observer A these two events occur at different positions (Figure 14.32*b*).

Relative to observer B, the throw and catch of particle 2 occur at the same position (Figure 14.32*c*), and so the time interval measured by B is the proper time interval. Because of symmetry, the collision seen by observer A is identical to that seen by observer B, and so the proper time interval measured by A between the throw and catch of particle 1 is equal to the proper time interval measured by B between the throw and catch of particle 2, as you saw in Checkpoint 14.17:

$$\Delta t_{A1} = \Delta t_{B2} = \Delta t_{\text{proper}}. \tag{14.36}$$

Equation 14.13, $\Delta t_v = \gamma \Delta t_{\text{proper}}$, tells us that if observer B measures the proper time interval between the throw and catch of particle 2, observer A moving relative to B obtains that the time interval between those events is $\gamma \Delta t_{\text{proper}}$, where γ depends on the relative velocity of A and B. Therefore

$$\Delta t_{A2} = \gamma \Delta t_{\text{proper}}, \tag{14.37}$$

where v in $\gamma = 1/\sqrt{1 - v^2/c_0^2}$ is the speed of B relative to A: $v = |\vec{v}_{A2x}|$. Substituting Eq. 14.36 into Eq. 14.37 yields $\Delta t_{A2} = \gamma \Delta t_{A1}$, and so Eq. 14.35 becomes

$$\Delta v_{A1y} = -\frac{d}{\Delta t_{A1}} \quad \text{and} \quad \Delta v_{A2y} = \frac{d}{\gamma \Delta t_{A1}}. \tag{14.38}$$

Because the two-particle system is isolated, the sum of the particles' momentum changes in reference frame A must add to zero:

$$\begin{aligned} \Delta p_{A1y} + \Delta p_{A2y} &= m_{v1} \Delta v_{A1y} + m_{v2} \Delta v_{A2y} \\ &= -m_{v1} \frac{d}{\Delta t_{A1}} + m_{v2} \frac{d}{\gamma \Delta t_{A1}} = 0, \end{aligned} \tag{14.39}$$

which means that in reference frame A,

$$m_{v2} = \gamma m_{v1}. \tag{14.40}$$

According to observer A, the inertia of particle 2 is greater than the inertia of particle 1.

14.18 (*a*) According to observer A in Figure 14.32, which particle is moving faster? (*b*) Does A conclude that the inertia of a particle increases or decreases with speed? (*c*) Answer these two questions for observer B.

QUANTITATIVE TOOLS

Equation 14.40 gives us the relationship between the inertias of two particles moving relative to each other and relative to us. To determine the relationship between the inertia of a particle moving relative to us and that of an identical particle at rest relative to us, consider the limit in which the y components of the velocities are very small. In that case, particle 1 in Figure 14.32*b* is virtually at rest in reference frame A, $v_{A1} \approx 0$, and particle 2 moves horizontally along with observer B at speed $v_{A2} \approx |\vec{v}_{A2x}| = v$. When the particle is at rest relative to us, its inertia and mass are equivalent and so $m_v \approx m$, where m is the mass of the particle. If we denote the inertia of the moving particle by m_v, Eq. 14.40 shows us how a particle's **inertia** m_v, which depends on speed, is related to its mass m, which is an invariant:

$$m_v = \gamma m. \tag{14.41}$$

The first thing to note is that for speeds up to about $0.1c_0$, $\gamma \approx 1$, and so mass and inertia are virtually indistinguishable, in agreement with what we established in Chapter 13. As the speed approaches c_0, however, m_v increases to infinity. Because inertia is a measure of an object's resistance to being accelerated, m_v approaching infinity means it becomes more and more difficult to speed up the particle as its speed approaches c_0. We therefore must conclude that no object or particle that has mass can be accelerated to speed c_0.

EXERCISE 14.10 Sluggish electron

By what factor does the inertia of an electron increase (*a*) as the electron is accelerated from rest to $\frac{1}{3}c_0$ and (*b*) as its speed is increased by another 20%? (*c*) Does the mass of the electron change as it is accelerated?

SOLUTION From Eq. 14.41, I know that the factor by which the inertia increases is $m_v/m = \gamma$. (*a*) At $\frac{1}{3}c_0$, $\gamma = 1/\sqrt{1-\left(\frac{1}{3}\right)^2} = 1.06$. The inertia increases by 6%. ✔

(*b*) Increasing the speed by another 20% means adding

$$\tfrac{1}{5}\left(\tfrac{1}{3}c_0\right) = \tfrac{1}{15}c_0$$

to the speed $\frac{1}{3}c_0$, giving

$$v = \tfrac{1}{3}c_0 + \tfrac{1}{15}c_0 = \tfrac{2}{5}c_0$$

$$\gamma = \frac{1}{\sqrt{1-\left(\frac{2}{5}\right)^2}} = 1.09.$$

The inertia of the electron increases by another 3%. ✔

(*c*) No. Mass is an invariant. Thus the mass of an object is independent of the motion of either the observer or the object. ✔

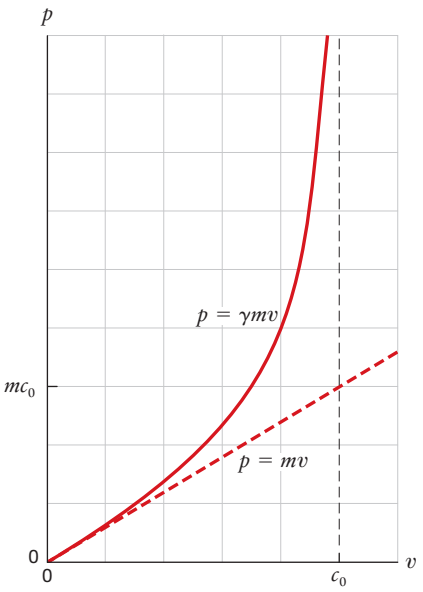
14.19 To what fraction of c_0 must an electron at rest in the Earth reference frame be accelerated in order for its inertia to increase by a factor of 1000?

Now that we have an expression for the inertia of a moving particle (Eq. 14.41), we can obtain an expression for the particle's momentum in terms of its mass and velocity. Substituting Eq. 14.41 into Eq. 14.34 gives

$$\vec{p} = \gamma m\vec{v}. \tag{14.42}$$

Unlike the expression for momentum you learned in Chapter 4, $\vec{p} = m\vec{v}$, Eq. 14.42 does not show the momentum increasing linearly with increasing velocity. At ordinary speeds, Eq. 14.42 reduces to the familiar form $\vec{p} = m\vec{v}$. As the speed approaches c_0, the momentum increases rapidly to infinity (**Figure 14.33**). Experiments such as the one described in the next example have conclusively shown that Eq. 14.42 correctly accounts for the momentum of a particle at all speeds up to c_0.

Figure 14.33 As the speed of a particle approaches c_0, the magnitude of its momentum approaches infinity. Up to a speed of about $0.1c_0$, the magnitude of the momentum is very close to mv.

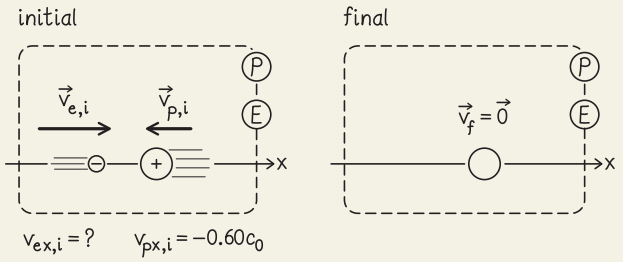


EXAMPLE 14.11 Fast collision

A fast-moving electron collides head-on with a proton moving at $0.60c_0$ as measured in the Earth reference frame. What is the electron's initial speed if the two particles combine to form a system that after the collision is at rest in the Earth reference frame? The mass of the proton is 1836 times that of the electron.

1 GETTING STARTED I begin by making a sketch showing the electron-proton system before and after the collision (**Figure 14.34**). I choose the positive x axis to be in the initial direction of motion of the electron. I can consider the electron-proton system to be isolated, and so the initial momentum of the system must be equal to the final momentum (which is zero).

Figure 14.34



2 DEVISE PLAN I can use Eq. 14.42 to calculate the initial momentum of each particle and then set the sum of these momenta equal to zero. The resulting expression contains four quantities: the unknown initial velocity component $v_{ex,i}$ of the electron, the initial velocity component $v_{px,i}$ of the proton ($-0.60c_0$), the mass m_p of the proton, and the mass m_e of the electron. Because I am given that $m_p/m_e = 1836$, I can solve for the unknown initial velocity of the electron.

3 EXECUTE PLAN The initial momentum of the system must be zero, so

$$p_{px,i} + p_{ex,i} = \gamma_{p,i} m_p v_{px,i} + \gamma_{e,i} m_e v_{ex,i} = 0, \qquad (1)$$

where I have added subscripts to the Lorentz factors to keep the factor for the proton separate from the factor for the electron. Dividing the two terms containing γ in Eq. 1 by m_e, and using Eq. 14.6 to write out the Lorentz factors, I can rewrite Eq. 1 as

$$\frac{m_p}{m_e}\frac{v_{px,i}}{\sqrt{1 - v_{px,i}^2/c_0^2}} = -\frac{v_{ex,i}}{\sqrt{1 - v_{ex,i}^2/c_0^2}}.$$

When I square both sides and then substitute the values for the mass ratio and the speed of the proton, I get

$$(1836)^2\frac{(-0.60c_0)^2}{1-(-0.60)^2} = \frac{v_{ex,i}^2}{1-(v_{ex,i}/c_0)^2},$$

which I can solve for $v_{ex,i}$:

$$v_{ex,i} = (1 + 5.3 \times 10^{-7})^{-1/2} c_0 = 0.99999974\ c_0. \checkmark$$

I show eight significant digits because the 1 in the above expression is exact (infinite number of significant digits) and so the sum $1 + 5.3 \times 10^{-7}$ is known to the eighth decimal place (1.00000053).

4 EVALUATE RESULT I note that this speed is close to c_0 (but not greater than or equal to it). The fact that the electron needs to travel at such a high speed is not surprising because the proton's mass is nearly 2000 times that of the electron.



14.20 In Example 14.11, (*a*) is the collision elastic, inelastic, or totally inelastic (see Section 5.1)? (*b*) How fast would the electron have to move if the expression for momentum were $\vec{p} = m\vec{v}$ instead of Eq. 14.42?

14.8 Conservation of energy

Now we turn our attention to energy. Suppose we accelerate a particle from rest to a speed approaching c_0.* To determine the increase in its kinetic energy, let us calculate the work done on the particle. From Chapter 9, we know that the work done on the particle is given by Eq. 9.22:

$$W = \int_{x_i}^{x_f} F(x)\,dx = \int_{x_i}^{x_f} \frac{dp}{dt}dx, \qquad (14.43)$$

where we have substituted the definition of force (Eq. 8.4). Substituting $p = \gamma mv$ and noting that the mass m is constant, we can write the integrand in Eq. 14.43 as

$$\frac{dp}{dt}dx = \frac{d}{dt}(\gamma mv)dx = m\frac{d(\gamma v)}{dt}dx. \qquad (14.44)$$

*From here on, speeds are assumed to be relative to the Earth reference frame unless stated otherwise.

QUANTITATIVE TOOLS

Note that because γ is a function of v, it varies with time and cannot be pulled out of the differentiation. Using $dx = v\,dt$ and canceling the dt factors, we get

$$\frac{dp}{dt}dx = m\frac{d(\gamma v)}{dt}dx = m\frac{d(\gamma v)}{dt}v\,dt = mv\,d(\gamma v), \qquad (14.45)$$

and so the work done on the particle is

$$W = m\int_{\text{i}}^{\text{f}} v\,d(\gamma v) = m\int_{\text{i}}^{\text{f}} (v^2\,d\gamma + v\gamma\,dv), \qquad (14.46)$$

where we have used the chain rule $d(\gamma v) = \gamma\,dv + v\,d\gamma$ (Appendix B).

The relationship between v^2 and γ is given by Eq. 14.9: $v^2/c_0^2 = 1 - 1/\gamma^2$. Differentiating both sides of Eq. 14.9 with respect to time yields

$$\frac{2v}{c_0^2}\frac{dv}{dt} = 0 - \frac{-2}{\gamma^3}\frac{d\gamma}{dt}, \qquad (14.47)$$

and then multiplying both sides by $\frac{1}{2}c_0^2\gamma\,dt$ yields

$$\gamma v\,dv = \frac{c_0^2}{\gamma^2}d\gamma. \qquad (14.48)$$

From Eq. 14.8, $1/\gamma^2 = 1 - v^2/c_0^2$, we can say

$$\frac{1}{\gamma^2} = \frac{c_0^2 - v^2}{c_0^2} \quad \text{and} \quad \frac{c_0^2}{\gamma^2} = c_0^2 - v^2, \qquad (14.49)$$

and so c_0^2/γ^2 in Eq. 14.48 can be replaced by $c_0^2 - v^2$, yielding $\gamma v\,dv = (c_0^2 - v^2)d\gamma$. Substituting this result into Eq. 14.46 and noting that $\gamma = 1$ when $v = 0$ at the lower integration boundary, we get

$$W = m\int_1^{\gamma} [v^2\,d\gamma + (c_0^2 - v^2)d\gamma] = mc_0^2\int_1^{\gamma} d\gamma = (\gamma - 1)mc_0^2. \qquad (14.50)$$

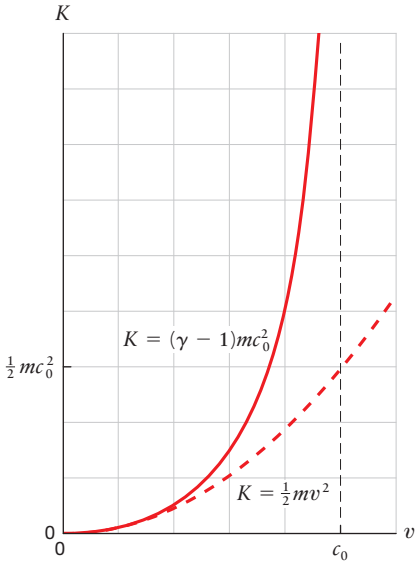
All of this work done on the particle goes toward increasing its kinetic energy, and so the particle's final **kinetic energy** is

$$K = (\gamma - 1)mc_0^2 = \gamma mc_0^2 - mc_0^2 = m_v c_0^2 - mc_0^2, \qquad (14.51)$$

where we have used the expression for the inertia, $m_v = \gamma m$, from Eq. 14.41.

This expression for the kinetic energy of a particle looks completely different from the expression we have used so far, $K = \frac{1}{2}mv^2$ (see Section 5.2). In particular, as the speed approaches c_0, the particle's kinetic energy becomes infinitely great (**Figure 14.35**). Therefore it would take an infinite amount of energy to accelerate a particle to speed c_0, so c_0 is the limiting speed at which any object can travel relative to another object.

Figure 14.35 As the speed of a particle approaches c_0, its kinetic energy approaches infinity. Up to speeds of about $0.25c_0$, the kinetic energy is very close to $\frac{1}{2}mv^2$.



14.21 Use Eq. 14.11, $\gamma - 1 \approx 0.5(v^2/c_0^2)$, to show that Eq. 14.51 reduces to the familiar $K = \frac{1}{2}mv^2$ at nonrelativistic speeds ($v \ll c_0$).

Each term on the right of Eq. 14.51 represents a quantity of energy, so Eq. 14.51 gives us the kinetic energy of a particle as the difference between two amounts of energy. In Chapter 5 we defined the energy of a system as the sum of that system's kinetic and internal energies, $E \equiv K + E_{\text{int}}$

(Eq. 5.21), and so we see that the kinetic energy of a system is obtained by subtracting the internal energy from the system's energy:

$$K = E - E_{\text{int}}. \tag{14.52}$$

In the right side of Eq. 14.51, note that the first term in the difference is proportional to inertia m_v and the second is proportional to mass m. In Section 14.4 we established that there is a relationship between inertia and energy and a relationship between mass and internal energy. Comparing Eqs. 14.51 and 14.52 thus suggests that the energy is given by

$$E = m_v c_0^2 \tag{14.53}$$

and the **internal energy** is given by

$$E_{\text{int}} = mc_0^2. \tag{14.54}$$

The assignment made in Eqs. 14.53 and 14.54 may look like a leap of faith, but in Section 14.4 we found that in a nuclear reaction between a proton and a deuteron, both the internal energy and the mass of the system decrease. Equation 14.54 tells us that the changes in internal energy and mass should be related as

$$\begin{aligned}\Delta E_{\text{int}} &= E_{\text{int,f}} - E_{\text{int,i}} \\ &= m_{\text{f}} c_0^2 - m_{\text{i}} c_0^2 = \Delta m c_0^2,\end{aligned} \tag{14.55}$$

or $\Delta E_{\text{int}}/\Delta m = c_0^2$. The experimental data show that the ratio $\Delta E_{\text{int}}/\Delta m = c_0^2$ is $8.98 \times 10^{16}\ \text{m}^2/\text{s}^2$ (see Example 14.5), which is precisely c_0^2, confirming Eq. 14.54. The term E_{int} in Eq. 14.54 thus represents the energy that corresponds to *all* of the mass of the particle.

All forms of internal energy we have discussed so far require an arbitrary choice of reference point (see, for example, Eq. 7.21), and so we usually considered only differences in internal energy. Equation 14.54, however, provides us with a measure of internal energy that does not depend on any choice of reference point. What is more, we can calculate E_{int} from the mass m, which is a quantity we can determine experimentally.

For example, a proton, as mentioned in Chapter 7, is made up of quarks held together by the strong nuclear interaction. If we add up the values of the masses of these quarks (which are determined from a combination of theory and experiment and known only approximately), we discover that their sum is just 1% of the mass of the proton. The remaining 99% of the mass of the proton must therefore come from the nuclear potential energy that binds the quarks together. Because protons (and neutrons, which also consist of three quarks) are the basic building blocks of all matter, it follows that most of what we consider matter really *is* energy! Even though we have no mathematical expression for nuclear potential energy, we can determine its value from the masses of the proton and the quarks, and the relationship $E_{\text{int}} = mc_0^2$.



14.22 (*a*) Draw a simplified energy diagram for the totally inelastic collision between two identical objects moving at speed v toward each other. After the collision, the two objects form a compound object that is at rest. (*b*) Which of the following quantities are constant in the collision: kinetic energy, mass, energy?

QUANTITATIVE TOOLS

EXAMPLE 14.12 Alpha emission

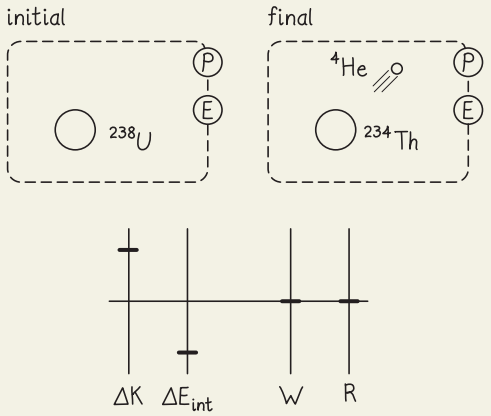
A uranium nucleus ^{238}U at rest decays to a thorium nucleus ^{234}Th by emitting an alpha particle (which is a helium nucleus ^{4}He):

$$^{238}\text{U} \rightarrow {}^{234}\text{Th} + {}^{4}\text{He}.$$

The mass of the uranium atom is $m_U = 395.292632 \times 10^{-27}$ kg, the mass of the thorium atom is $m_{Th} = 388.638541 \times 10^{-27}$ kg, and the mass of the helium atom is $m_{He} = 6.646479 \times 10^{-27}$ kg. What is the combined kinetic energy of the fragments?

1 GETTING STARTED I begin by drawing a simplified energy diagram for the alpha emission (**Figure 14.36**), choosing the uranium nucleus and the fragments as my (closed) system. Because the system is closed, the change in kinetic energy is the negative of the change in internal energy: $\Delta K = -\Delta E_{int}$.

Figure 14.36



2 DEVISE PLAN Because the initial kinetic energy is zero, the final kinetic energy of the fragments is equal to $-\Delta E_{int}$. To determine this change in internal energy, I can calculate the change in the mass of the system and then use Eq. 14.55.

3 EXECUTE PLAN

$$\Delta m = (m_{Th} + m_{He}) - m_U = -7.61 \times 10^{-30} \text{ kg}.$$

Substituting this value into Eq. 14.55 yields

$$\Delta E_{int} = \Delta m c_0^2 = (-7.61 \times 10^{-30} \text{ kg})(3.00 \times 10^8 \text{ m/s})^2$$

$$= -6.85 \times 10^{-13} \text{ J},$$

so the final kinetic energy of the fragments is 6.85×10^{-13} J. ✔

4 EVALUATE RESULT The answer I obtain is a tiny amount of energy, contrary to my expectations for nuclear energy! However, this is the energy liberated by the nucleus of a single atom. One atom is about 10^{-10} m in diameter, which means a cube of matter 1 mm on a side contains about $[(10^{-3} \text{ m})/(10^{-10} \text{ m})]^3 = 10^{21}$ atoms. If each atom liberates an amount of energy equal to the answer I got, my cube yields $(10^{21})(6.85 \times 10^{-13} \text{ J}) = 7 \times 10^8$ J of energy, which is a very large amount and more in line with what I'd expect.



14.23 An atom in the filament of a light bulb emits energy in the form of light. (*a*) Draw a simplified energy diagram representing this emission. (*b*) Does the mass of the atom increase, decrease, or stay the same? (*c*) Does the inertia of the atom increase, decrease, or stay the same? Assume the atom remains at rest in the Earth reference frame.

We can obtain a useful relationship from our expressions for momentum, $\vec{p} = \gamma m \vec{v}$ (Eq. 14.42), and energy, $E = m_v c_0^2$ (Eq. 14.53), by evaluating the quantity $E^2 - (c_0 p)^2$. Substituting γm for m_v in Eq. 14.53, we have

$$E^2 - (c_0 p)^2 = \gamma^2 m^2 c_0^4 - c_0^2 \gamma^2 m^2 v^2 = \gamma^2 m^2 c_0^2 (c_0^2 - v^2). \quad (14.56)$$

From Eq. 14.49, we see that $c_0^2 - v^2 = c_0^2/\gamma^2$, and so

$$E^2 - (c_0 p)^2 = (m c_0^2)^2. \quad (14.57)$$

The right side of this expression is an invariant, and so its value does not depend on the reference frame. The left side, however, contains the energy E and the momentum magnitude p of the system under consideration, both of which depend on the reference frame! Again, we have constructed an invariant from two quantities that are not invariants themselves. Equation 14.57 tells us that if we measure the energy and the magnitude of the momentum of a system in a certain reference frame and then compute the quantity $E^2 - (c_0 p)^2$, we obtain the same value as the corresponding

combination in any other reference frame. In each case, the result is equal to $(mc_0^2)^2$, and so the quantity $E^2 - (c_0p)^2$ is an invariant.

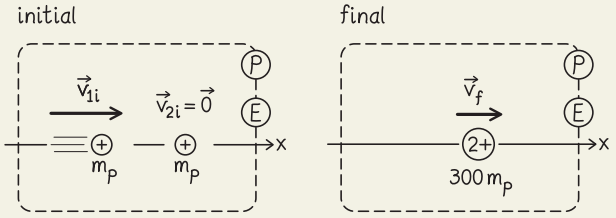
Energy and momentum are conserved quantities when evaluated with respect to each reference frame, and so for a system that is both isolated and closed, the energy and momentum laws tell us that $E_i = E_f$ and $\vec{p}_i = \vec{p}_f$. We have used these equations extensively to determine E_f and p_f given E_i and p_i. Equation 14.57 relates values of E and p to values of m. Because values of m are known from separate measurements, Eq. 14.57 allows us to eliminate momentum from the problem. The next example and checkpoint illustrate the usefulness of Eq. 14.57.

EXAMPLE 14.13 Proton collision

Two protons, each of mass $m_p = 1.67 \times 10^{-27}$ kg, collide to produce a particle that has a mass 300 times the mass of each proton. What is the minimum kinetic energy that one of the protons must have in order to create this particle if the other proton is initially at rest?

1 GETTING STARTED I begin by organizing the information in a sketch (**Figure 14.37**). I denote the incoming proton by 1 and the proton initially at rest by 2. The two protons form an isolated, closed system, and so the energy law requires $E_{1i} + E_{2i} = E_f$ and the momentum law requires $\vec{p}_{1i} = \vec{p}_f$ (because $\vec{p}_{2i} = 0$).

Figure 14.37



2 DEVISE PLAN The initial kinetic energy K_{1i} of proton 1 is given by Eq. 14.52, with the internal energy E_{1int} given by Eq. 14.54. I must therefore determine the initial energy E_{1i} of proton 1. I can use Eq. 14.57 to relate the energy E of each particle to its momentum p and then combine this with the energy and momentum law to solve for the initial energy of proton 1.

3 EXECUTE PLAN Applying Eq. 14.57 to proton 1, proton 2, and the final particle, I get

$$E_{1i}^2 - (c_0p_{1i})^2 = (m_pc_0^2)^2 \quad (1)$$

$$E_{2i}^2 = (m_pc_0^2)^2 \quad (2)$$

$$E_f^2 - (c_0p_f)^2 = (300m_pc_0^2)^2. \quad (3)$$

Substituting $E_{1i} + E_{2i} = E_f$ and $p_{1i} = p_f$ into Eq. 3 yields

$$(E_{1i} + E_{2i})^2 - (c_0p_{1i})^2 = (300m_pc_0^2)^2.$$

Working out the left side, I get

$$E_{1i}^2 + 2E_{1i}E_{2i} + E_{2i}^2 - (c_0p_{1i})^2 = E_{1i}^2 - (c_0p_{1i})^2 + E_{2i}^2 + 2E_{1i}E_{2i}$$

$$= (300m_pc_0^2)^2,$$

so, after substituting from Eqs. 1 and 2, I have

$$(m_pc_0^2)^2 + (m_pc_0^2)^2 + 2E_{1i}(m_pc_0^2) = (300m_pc_0^2)^2.$$

Solving for E_{1i} gives

$$E_{1i} = \frac{(300m_pc_0^2)^2 - 2(m_pc_0^2)^2}{2(m_pc_0^2)} = \tfrac{1}{2}(300^2 - 2)m_pc_0^2. \quad (4)$$

From Eqs. 14.52 and 14.54, I have $K_{1i} = E_{1i} - m_pc_0^2$, so

$$K_{1i} = \left[\tfrac{1}{2}(300^2 - 2)m_pc_0^2\right] - m_pc_0^2 = \tfrac{1}{2}(300^2 - 4)m_pc_0^2$$

$$= \tfrac{1}{2}(300^2 - 4)(1.67 \times 10^{-27}\text{ kg})(3.00 \times 10^8\text{ m/s})^2$$

$$= 6.76 \times 10^{-6}\text{ J}. \checkmark$$

4 EVALUATE RESULT My answer is large for a particle that has a mass on the order of 10^{-27} kg. Comparing my result in Eq. 4 with Eq. 14.51, however, I see that $\gamma - 1 = (300^2 - 4)/2$, or $\gamma = 4.5 \times 10^4$, and so the proton has to be accelerated to very close to the speed of light. It is therefore not unreasonable that the energy of the proton is on the order of microjoules.

14.24 In Example 14.13, what is the minimum kinetic energy of each proton if they have the same initial speed and collide head-on?

QUANTITATIVE TOOLS

Chapter Summary

Time measurements (Sections 14.1, 14.2)

Concepts An **event** is a physical occurrence that takes place at a specific location at a specific instant.

Two events are **simultaneous** when they are observed to occur at the same instant in a reference frame made up of a grid of synchronized clocks. Events that are simultaneous in one reference frame need not be simultaneous in another reference frame moving relative to the first one.

An **invariant** is a quantity whose value does not depend on the choice of reference frame.

The **speed of light** is independent of the motion of the light source relative to any moving observer. Therefore the speed of light is an invariant. It is the limiting speed of any object moving relative to any other object.

Quantitative Tools In SI units, the speed of light in vacuum is $c_0 = 3.00 \times 10^8$ m/s.

Space-time (Sections 14.3, 14.5, 14.6)

Concepts The **proper time interval,** Δt_{proper}, is the time interval between two events that occur at the same position in a particular reference frame but at different instants. An observer measuring the proper time interval between two events records the shortest time interval between those events.

Time dilation: A clock moving relative to an observer is measured to be running more slowly than an identical clock at rest relative to that observer.

The **proper length,** ℓ_{proper}, of an object is the length of that object measured by an observer who is at rest relative to the object.

Length contraction: When an object moves at speed v relative to an observer, the observer measures the length ℓ_v of the object along the direction of its motion to be shorter than the proper length of the object.

Quantitative Tools For an object moving at speed v, the **Lorentz factor** γ is a measure of the extent to which relativity affects our everyday notions of length and time intervals:

$$\gamma \equiv \frac{1}{\sqrt{1 - v^2/c_0^2}}. \tag{14.6}$$

At nonrelativistic speeds ($v \leq 0.1c_0$), $\gamma \approx 1$, but it approaches infinity as v approaches c_0.

Time dilation: If Δt_{proper} is the proper time interval between two events measured by an observer for whom the two events occur at the same position but at different instants, then the time interval Δt_v between those events measured by an observer moving at speed v relative to the observer is

$$\Delta t_v = \gamma \Delta t_{\text{proper}}. \tag{14.13}$$

The **space-time interval** s^2 between two events that are separated by a time interval Δt and a space interval Δx is given by

$$s^2 \equiv (c_0 \Delta t)^2 - (\Delta x)^2. \tag{14.18}$$

If an object that has proper length ℓ_{proper} moves at speed v relative to an observer, that observer measures the length ℓ_v of the object to be

$$\ell_v = \frac{\ell_{\text{proper}}}{\gamma}. \tag{14.28}$$

The **Lorentz transformation equations** are relationships between the coordinates of an event e measured in two reference frames A and B that coincide at $t_A = t_B = 0$ and are moving in the x direction at constant velocity $\vec{v}_{AB}$ relative to each other:

$$t_{Be} = \gamma\left(t_{Ae} - \frac{v_{ABx}}{c_0^2} x_{Ae}\right) \tag{14.29}$$

$$x_{Be} = \gamma(x_{Ae} - v_{ABx} t_{Ae}) \tag{14.30}$$

$$y_{Be} = y_{Ae} \tag{14.31}$$

$$z_{Be} = z_{Ae}. \tag{14.32}$$

Velocity addition: From the Lorentz transformation equations, we find that v_{Aox}, the x component of the velocity of an object o measured in reference frame A, is related to v_{Box}, the x component of its velocity measured in reference frame B, by

$$v_{Box} = \frac{v_{Aox} - v_{ABx}}{1 - \frac{v_{ABx}}{c_0^2} v_{Aox}}. \tag{14.33}$$

Energy and momentum (Sections 14.4, 14.7, 14.8)

Concepts The internal energy E_{int} of an object (energy not associated with the object's motion) is an invariant and is proportional to the object's mass m (the property that determines its gravitational interaction). Mass is also an invariant.

Unlike mass, inertia (a quantitative measure of an object's tendency to resist any change in its velocity) is not an invariant: What an observer measures for an object's inertia depends on the motion of the observer relative to the object.

Quantitative Tools For an object of mass m, its inertia m_v is

$$m_v = \gamma m, \tag{14.41}$$

its momentum $\vec{p}$ is

$$\vec{p} = \gamma m \vec{v}, \tag{14.42}$$

and its kinetic energy K is

$$K = (\gamma - 1) m c_0^2. \tag{14.51}$$

The energy E of an object of mass m and inertia m_v is

$$E = m_v c_0^2, \tag{14.53}$$

and its internal energy E_{int} is

$$E_{int} = m c_0^2. \tag{14.54}$$

Questions and Problems

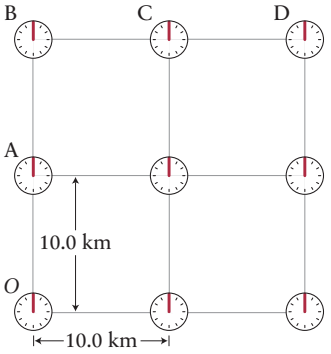
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

14.1 Time measurements

1. Which of the following are events? (*a*) The local television station's newscast runs from 11:00 to 11:30 p.m. (*b*) The ceremonial first pitch is thrown at a baseball game in Omaha, Nebraska. (*c*) The Perseid meteor shower is seen across the eastern coast of North America. (*d*) A rainstorm lasts all day in Kyoto, Japan. (*e*) Two cars crash in front of your home. (*f*) An all-night concert starts in a local park. (*g*) Light emitted by a supernova reaches a spherical shell of radius 40,000 light years relative to its source. (One light year is the distance light travels in one year.) •
2. You are in a rocket moving away from Earth at one-third the speed of light relative to Earth. A friend is on Earth, and an astronaut in another rocket is moving toward Earth at one-third the speed of light (in the Earth reference frame), on a path collinear with your path. If each of you records the duration of your journey in your three respective reference frames, which of you records the proper time interval? •
3. A set of atomic clocks is placed on a square grid as shown in Figure P14.3. In order to synchronize these clocks, you set the time on each to a certain reading before they are allowed to run. Clock O starts first, and at precisely 12:00 noon, it sends out a light pulse in all directions. Each of the other clocks starts running as soon as the pulse reaches it. What initial time should you use for clocks A, B, C, and D to synchronize them with clock O? Express your answer to the nearest microsecond before or after 12:00 noon. ••

Figure P14.3



4. Events 1 and 2 are exploding firecrackers that each emit light pulses. In the reference frame of the detector, event 1 leaves a char mark at a distance 3.40 m from the detector, and event 2 leaves a similar mark at a distance 2.10 m from the detector. If the two events are simultaneous in the reference frame of the detector and occur at instant $t = 0$, at what instant of time will each light pulse be detected? ••
5. Consider the grid of clocks in Figure 14.6. Do observers stationed at clocks far from the reference clock read earlier, later, or the same time as the time on the reference clock (*a*) if the clocks are to be synchronized after a pulse is sent out activating them but this pulse has not yet been sent, and (*b*) if the pulse has been sent out and all clocks are running? (*c*) Suppose the pulse has been sent and all clocks are running. If you stand at the reference clock and use a telescope to read the time on distant clocks, do you read a time earlier, later, or the same as the time on the reference clock? ••
6. Observers A and B are both at rest in the Earth reference frame, in different parts of a large city, awaiting the start of a fireworks show. They want to synchronize their atomic clocks for some experiments to be conducted later that evening. A and B plan to start their clock displays at an agreed signal, so each clock display is preset to exactly 10:00 p.m., but they will not record time until the "start" buttons are pressed. Observer A suggests that each of them start upon *seeing* the first firework, and observer B suggests that each of them start upon *hearing* the first firework. (*a*) Which plan leads to the clocks being more closely synchronized, and why? (*b*) Is it possible for the two methods to give the same agreement between clocks? •••

14.2 *Simultaneous* is a relative term

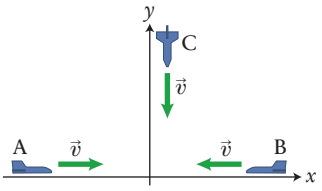
7. Observer A is stationary in the Earth reference frame, and pilot B is flying toward pilot A at relative speed v_{AB}. What speed does observer A measure for the speed of the light emitted by a signal light on pilot B's ship? •
8. Alarm clock 1 is positioned at $x = -d$ and alarm clock 2 at $x = +d$ in reference frame A. According to a clock in reference frame A, both go off at $t = 0$. An observer in reference frame B, which is moving relative to A, measures alarm clock 2 as going off first. In what direction is B moving relative to A? •
9. You are traveling toward a large fixed mirror at a constant relative speed $0.20c_0$. At $t = 0$, when the mirror is a distance d from you (measured in your reference frame), you emit a light pulse from a lantern and then detect the reflected light pulse 0.80 μs later. What is the value of d? ••
10. You are in a jet helicopter traveling horizontally at 180 m/s relative to Earth. At one instant you observe the light pulses from two lightning strikes, one directly ahead of the helicopter and one directly behind it. (*a*) Can you say that the strikes occurred simultaneously in your reference frame? (*b*) How would your answer change if you were instead hovering at rest relative to the ground when you observed the two pulses? (*c*) How would your answer change if the pilot told you that the two strikes occurred at equal distances from the helicopter in the helicopter's reference frame? (*d*) Would an observer on the ground agree with your answers to parts *a*, *b*, and *c*? ••
11. Consider a lightweight, straight, rigid rod that has a proper length of 1×10^6 m. If you rotate the rod at 100 turns/s about a perpendicular axis through its center, at what speed do the ends of the rod move? What prevents such a case from occurring? ••
12. Observer A on Earth sees spaceship B moving at speed $v = 0.600c_0$ away from Earth and spaceship C moving at the same speed in the opposite direction. Observer A determines that both ships simultaneously sent out a radio signal at 12:00:00 noon and another signal at 12:05:00 p.m., making the time interval A measures between the emission of the signals from each ship $\Delta t_A = 5.00$ min. Let Δt_{BB} denote the time interval measured on ship B between the emission of signals on ship B, Δt_{BC} denote the time interval measured on ship B between the emission of signals on ship C, and Δt_{CB} denote the time interval measured on ship C between emission of signals on ship B. (*a*) Is Δt_{BB} greater than, less than, or equal to Δt_A? (*b*) Is Δt_{BC} greater than, less than, or equal to Δt_{CB}? ••

14.3 Space-time

13. An astronaut takes what he measures to be a 10-min nap in a space station orbiting Earth at 8000 m/s. A signal is sent from the station to Earth at the instant he falls asleep, and a second signal is sent to Earth the instant he wakes. Does an observer on Earth measure the astronaut's nap as being shorter than, longer than, or equal to 10 min? •
14. Suppose you could move at nearly the speed of light. If you were to move at this speed for a large portion of your life, would it be possible for you to live long enough to see a later calendar year on Earth than you otherwise would have? Compare your answer with that for Checkpoint 14.8. •
15. Standing somewhere between two vertical mirrors, you hold a lantern and at $t = 0$ emit a light pulse that travels in all directions. You observe the pulse reflected from the mirror on your right at $t = 2.5$ μs and from the mirror on your left at $t = 6.5$ μs. What is the distance between the two mirrors in your reference frame? •
16. A spaceship of proper length $\ell = 100$ m travels in the positive x direction at a speed of $0.800c_0$ relative to Earth. An identical spaceship travels in the negative x direction along a parallel course at the same speed relative to Earth. At $t = 0$, an observer on Earth measures a distance $d = 50{,}000$ km separating the two ships. At what instant does this observer see the leading edges of the two ships pass each other? ••
17. A spacecraft traveling away from Earth at relative speed $0.850c_0$ sends a radio message when it is 65,000,000 km away from Earth in the Earth reference frame. When the signal is received, Earth-based engineers immediately send a reply to the spacecraft. What do these engineers measure for the time interval from the instant the spacecraft sends the original message to the instant the spacecraft receives the reply? ••

18. Spaceships A, B, and C in Figure P14.18 all have the same proper length and all fly with the same speed according to an observer on Earth. Rank the lengths of the three ships as measured by an observer in ship A, longest first. ●●

Figure P14.18



19. The boundary of a lunar base is a square 1 km wide and 1 km long in the Moon reference frame. Spaceship A flies at high relative speed close to the Moon parallel to the edge that we shall call the length of the field. Crew A measures the length and width, and reports that the area of the field is 0.5 km^2 in reference frame A. Spaceship B flies with the same relative speed as spaceship A, but parallel to one diagonal of the field. Crew B measures both diagonals. What surface area does crew B calculate in reference frame B? ●●●

14.4 Matter and energy

20. What is the kinetic energy of a pin in your reference frame if you measure its inertia to be three times its mass? ●
21. A golf ball hit from a tee accelerates from rest to a speed of 40.0 m/s relative to the Earth reference frame. By what percent does the mass of the ball increase as it is hit? ●
22. One end of a vertical spring of spring constant $k = 1500$ N/m is attached to the floor. You compress the spring so that it is 2.40 m shorter than its relaxed length, place a 1.00-kg ball on top of the free end, and then release the system at $t = 0$. (All values are measured in the Earth reference frame.) (*a*) By how much does the mass of the spring change during the time interval from $t = 0$ to the instant the ball leaves the spring? (*b*) By how much does the mass of the Earth-ball-spring system change during the time interval from $t = 0$ to the instant the ball reaches its maximum height? ●●
23. You pick up your physics book from the floor and put it on your desk. In the Earth reference frame, which of the following quantities have changed for the system made up of Earth and the book: mass, energy, inertia, kinetic energy? ●●
24. Observer A is on Earth looking at the Moon, and observer B is on the Moon looking at Earth. Which of the following quantities do the observers agree on for the Earth-Moon system: inertia, energy, mass, kinetic energy? ●●
25. How far above Earth's surface must a 10,000-kg boulder be moved to increase the mass of the Earth-boulder system by 2.50 mg? Assume the same ratio of energy change to mass change as in Exercise 14.5(*b*): 8.98×10^{16} J/kg. ●●

14.5 Time dilation

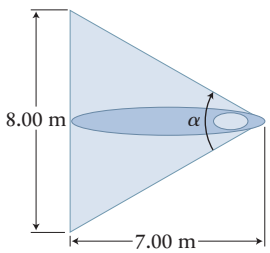
26. A bullet is fired from a rifle (event 1) and then strikes a soda bottle, shattering it (event 2). Is there some inertial reference frame in which event 2 precedes event 1? If so, does the existence of this reference frame violate causality? ●
27. Classify the space-time interval between each pair of events as timelike, spacelike, or lightlike: (*a*) A sunspot erupts on the surface of the Sun and is seen on Earth 500 s later. (*b*) A supernova explodes and is noticed 8.00×10^{20} m away after 8.45×10^4 y by an observer at rest relative to the supernova. (*c*) You receive light in the Earth reference frame emitted by a satellite 60,000 km away at $t = 0$, and 233 ms later receive light emitted by a spaceship 180,000 km away in the same direction as the satellite. (*d*) At $t = 0$, a pilot traveling at $0.75c_0$ relative to a star system receives signals from two planets A and B in that system. She then passes planet A (event 1) at $t = 450$ s and 90 s after that passes planet B (event 2). ●
28. Calculate the Lorentz factor for an object moving relative to Earth at (*a*) 60 mi/h, (*b*) 0.34 km/s (speed of sound in air), (*c*) 28×10^3 km/h (orbital speed of the International Space Station), (*d*) $0.1c_0$, (*e*) $0.3c_0$, and (*f*) $0.995c_0$. Express your answers as "1 + correction." ●●
29. You measure the period of light clock A to be T as it moves relative to you with a speed $0 < v_A < 0.5c_0$. You measure the period of an identical light clock (clock B) to be $3T$ as it moves relative to you with speed $v_B = 2v_A$. Determine numerical values for the speeds of each clock in your reference frame. ●●
30. A space station sounds an alert signal at time intervals of 1.00 h. Spaceships A and B pass the station, both moving at $0.400c_0$ relative to the station but in opposite directions. How long is the time interval between signals (*a*) according to an observer on A and (*b*) according to an observer on B? (*c*) At what speed must A move relative to the station in order to measure a time interval of 2.00 h between signals? ●●
31. A certain muon detector counts 600 muons per hour at an altitude of 1900 m and 380 muons per hour at sea level. Given that the muon half-life at rest is 1.5×10^{-6} s, determine the speed of the muons relative to Earth, assuming they all have the same speed. ●●
32. Planets A and B are 10 light years apart in the reference frame of planet A. (One light year is the distance light travels in one year.) A deep-space probe is launched from A, and 5 y later (in reference frame A) a similar probe is launched from B. Does a reference frame exist in which these two events (*a*) are simultaneous and (*b*) occur at the same position? (*c*) Compute, if possible, the proper time interval between the events? (*d*) How do your answers to parts *a*–*c* change if the planets are (in reference frame A) 5 light years apart and the time interval between launch from A and launch from B is 10 y? ●●
33. Section 14.3 describes how the number of muons reaching Earth's surface is greater than the number expected based on the muon half-life of 1.5×10^{-6} s. How fast, relative to Earth, must muons be moving in order for one of every million muons to pass through a distance equal to Earth's diameter without decaying? ●●

34. One cosmonaut orbited Earth for 437 days, as measured by Earth clocks. His speed of orbit was 7700 m/s relative to Earth during this time interval. Assume two clocks were synchronized on Earth, and one went into space with the cosmonaut while the other remained on Earth. You are interested in knowing by how much the clock readings disagree at the end of the 437 days. What are some problems with applying Eq. 14.13 to this situation? What information other than the cosmonaut's speed relative to Earth might you need to know? If you used Eq. 14.13, would you expect the clock readings to be very different from each other? •••

14.6 Length contraction

35. A cosmic ray traveling at $0.400c_0$ relative to Earth passes observer A (event 1) and then a short time interval later passes observer B (event 2). The observers are at rest, 8.00 km apart in the Earth reference frame. What are (*a*) the proper length and (*b*) the proper time interval between these events? •
36. An elementary particle is launched from Earth toward the Regulus system, 77.5 light years distant. At what speed relative to Earth must this particle travel to make this trip in 10 y in the reference frame of the particle? •
37. Spaceships A and B pass a space station at the same instant, the two ships moving in opposite directions at the same speed of $0.600c_0$ relative to the station. What is the Lorentz factor associated with the relative motion of the ships? •
38. Two elementary particles pass each other moving in opposite directions, each moving at speed $0.800c_0$ relative to Earth. What is the speed of one particle relative to the other? •
39. The radius of Earth is 6370 km in the Earth reference frame. In the reference frame of a cosmic ray moving at $0.800c_0$ relative to Earth, how wide does Earth seem (*a*) along the flight direction and (*b*) perpendicular to the flight direction? •
40. Show that the Lorentz transformation equations (Eqs. 14.29–14.32) reduce to the Galilean transformation equations in the limit $v_{AB} \ll c_0$. ••
41. An elementary particle moving away from Earth reaches its destination after two weeks (in the reference frame of the particle) traveling at $0.9990c_0$ relative to Earth. How far has it traveled in (*a*) its reference frame and (*b*) the Earth reference frame? ••
42. Describe the shape of the Moon as measured by an observer in a reference frame traveling past the Moon at a relative speed of (*a*) 1000 m/s, (*b*) $0.50c_0$, and (*c*) $0.95c_0$. ••
43. When at rest in the Earth reference frame, the delta-wing spaceship in Figure P14.43 is 7.00 m long and has a wingspan of 8.00 m. (*a*) What is the opening angle α of its wings? (*b*) If the same ship is moving forward at $0.700c_0$ relative to Earth, what does an observer on Earth measure for the plane length, wingspan, and opening angle α? ••

Figure P14.43



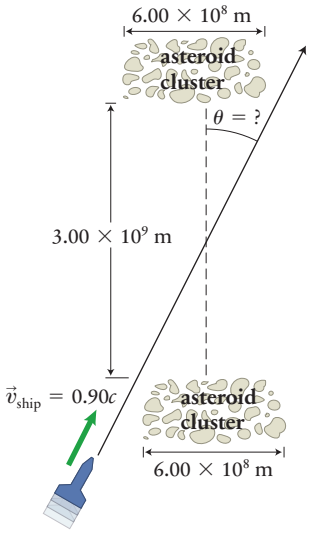
44. Standing at rest in the Earth reference frame, you observe two events, also at rest in that reference frame, occurring at these coordinates: event 1: $x_1 = 10$ km, $y_1 = 5.0$ km, $z_1 = 0$, $t_1 = 20\ \mu$s; event 2: $x_2 = 30$ km, $y_2 = 5.0$ km, $z_2 = 0$, $t_2 = 30\ \mu$s. At what speed and in what direction must your friend travel relative to Earth so as to observe these events as being simultaneous? ••
45. Use the Lorentz transformation equations (Eq. 14.29–14.32) to prove that the space-time interval s^2 between two events is *Lorentz-invariant*, which means it has the same value for all possible inertial reference frames. ••
46. An elementary particle travels at $0.840c_0$ across a solar system that has a diameter of 8.14×10^{12} m (both measurements in the reference frame S of the solar system). (*a*) What value does an observer in reference frame S measure for the time interval needed for the particle to cross the solar system? (*b*) What time interval is needed in the reference frame P of the particle? (*c*) What is the proper length of the diameter? (*d*) What is the proper time interval for the particle's trip across the system? (*e*) What is the diameter of the solar system in reference frame P? ••
47. A space colony defines the origin of a coordinate system in reference frame C, and you are at rest in this reference frame at position $x = +3.000 \times 10^5$ km. A ship passes the colony moving in the positive x direction at $0.482c_0$ in reference frame C. By prearrangement, observers on the ship (reference frame S) and on the colony use this event (event 0) to define $x = 0$ and $t = 0$ in both reference frames. Just as the ship passes, the colonists send out a light pulse toward you (event 1), and you transmit a reply pulse (event 2) immediately upon receipt. (*a*) At what position in reference frame S do observers see the colonists' pulse originate? (*b*) At what instant does the pulse originate according to these observers? (*c*) At what position in reference frame S do the observers see your reply pulse originate? (*d*) At what instant does the reply pulse originate according to observers on the ship? ••
48. Observer O at the origin of a coordinate system is at rest relative to two equidistant space stations located at $x = +3.00 \times 10^6$ km (station A) and $x = -3.00 \times 10^6$ km (station B) on the x axis. In reference frame O, station A sends out a light pulse at $t = 0$ (event 1) and station B also sends out a light pulse at $t = 0$ (event 2). Observer C moves relative to O with velocity $0.600c_0$ in the positive x direction, and observer D moves relative to O with velocity $0.600c_0$ in the negative x direction. What are the displacement from event 1 to event 2, and the time interval between events 1 and 2, according to (*a*) observer C and (*b*) observer D? ••
49. A spaceship travels at $0.610c_0$ away from Earth. When the ship is 3.47×10^{11} m away in the Earth reference frame, the ship clock is set to $t = 0$ and a message (light signal) is sent to its ground crew on Earth. When this message is received, the ground crew immediately sends a reply (light signal) to the ship. The ground crew adjust their clocks to specify $t = 0$ as the time the message was sent from the ship. (*a*) What is the reading on the ground crew's clock at the instant the ship receives the reply signal? (*b*) According to the ground crew, how far from

Earth is the ship when it receives the reply? (*c*) What is the reading on the ship's clock at the instant the reply reaches the ship? •••

50. According to an observer at rest in the Earth reference frame, ship A travels in one direction at speed $0.862c_0$ while ship B moves in the opposite direction at $0.717c_0$. Ship A sends out a shuttle module M that travels to ship B at speed $0.525c_0$ relative to A. For every 1.00 s that passes on M, how many seconds pass according to (*a*) a clock on A, (*b*) a clock on B, and (*c*) a clock held by the observer on Earth? •••

51. Two asteroid clusters at rest with respect to each other are 3.00×10^9 m apart, and both measure 6.00×10^8 m across in the reference frame of the clusters. A pilot in a spaceship moving at $0.900c_0$ relative to the cluster reference frame is trying to pass between the two clusters. To an observer who is stationary in the reference frame of the clusters, the ship just barely misses the corners of the clusters as shown in Figure P14.51. If a line were drawn between the centers of the two clusters, what would be the angle between that line and the ship's path in (*a*) the reference frame of the clusters and (*b*) the reference frame of the ship? •••

Figure P14.51



14.7 Conservation of momentum

52. Determine the magnitude of the momentum of a muon in a reference frame in which the muon moves with speed $0.500c_0$. The mass of a muon is 207 times the mass of an electron. •

53. An electron accelerates from $0.700c_0$ to $0.900c_0$ in the Earth reference frame. In this reference frame, by what factor do (*a*) its mass and (*b*) its inertia increase? •

54. (*a*) What is the magnitude of the momentum of an electron in a reference frame in which it is moving at $0.500c_0$? (*b*) At what speed must the electron move in a reference frame in which it has twice the momentum you calculated in part *a*? ••

55. Object A, mass 4.24×10^5 kg, is at rest in reference frame A, and object B, mass 7.71×10^4 kg, moves at $0.875c_0$ in this reference frame. (*a*) What does an observer at rest in A measure for the momentum of B? (*b*) What does an observer at rest with respect to object B measure for the momentum of A? (*c*) A reference frame C moves at $0.300c_0$ relative to reference frame A and moves in the same direction as object B. What are the momenta of objects A and B as measured by an observer at rest in reference frame C? ••

56. A uniform 200-kg cube that has a volume of 8.00 m^3 (measured when the cube is at rest in the Earth reference frame) travels perpendicular to a pair of its faces at $0.672c_0$ relative to Earth. (*a*) What is the mass density of the cube according to an observer at rest on Earth? (*b*) How fast (in the Earth reference frame) must a 100-kg object travel in the opposite direction to stop this cube in a collision? ••

57. Muons are formed 10.0 km above Earth's surface. What magnitude of momentum (in the Earth reference frame) must they have immediately after formation in order for only one of every 10,000 of them to reach the ground? The mass of a muon is 207 times the mass of an electron. ••

58. Two chunks of rock, each having a mass of 1.00 kg, collide in space. Just before the collision, an observer at rest in the reference frame of a nearby star determines that rock A is moving toward the star at $0.500c_0$ and rock B is moving away from the star at $0.300c_0$. If the rocks stick together after the collision, what are the direction and magnitude of their velocity (in the reference frame of the star) immediately after the collision? ••

59. In the Earth reference frame, a 150-kg probe travels at $0.860c_0$ while a 250-kg probe travels at $0.355c_0$ in the opposite direction. What are the velocities of the probes in the zero-momentum reference frame? •••

14.8 Conservation of energy

60. At what speed must a particle move in your reference frame so that its kinetic energy is equal to its internal energy? •

61. Particles A, B, and C each have mass *m*, and their energies are $E_A = E$, $E_B = 2E$, and $E_C = 3E$. Rank the particles, largest first, in order of (*a*) Lorentz factor, (*b*) kinetic energy, (*c*) speed, and (*d*) magnitude of momentum. •

62. What is the internal energy of an electron moving at $0.750c_0$ in the Earth reference frame? •

63. Uranium-238 decays to thorium and helium according to the reaction $^{238}\text{U} \rightarrow {}^{234}\text{Th} + {}^{4}\text{He}$. How much energy is released when 1.00 kg of uranium-238 decays? Compare this amount with the energy released when 1.00 kg of coal is burned (approximately 30 MJ). The atomic masses are $m_U = 395.292632 \times 10^{-27}$ kg, $m_{Th} = 388.638541 \times 10^{-27}$ kg, and $m_{He} = 6.646479 \times 10^{-27}$ kg. ••

64. Assume that 437 days is a reasonable limit for how long a human can endure constant-velocity space travel. Proxima Centauri, the star closest to our Sun, is 4.24 light years away from Earth. If you wanted to fly to Proxima Centauri within the 437-day limit in a rocket of mass 2.00×10^6 kg, how much energy would be required to accelerate the rocket to the necessary speed in the Earth reference frame? For this rough estimate, ignore the energy required to stop, and disregard the hundreds of days required at each end of the journey for acceleration.

Compare your answer with 0.6×10^{21} J, the yearly worldwide energy consumption predicted for 2015. ••

65. A particle of mass m_{orig}, initially at rest in the Earth reference frame, decays into two particles 1 and 2 that have masses m_1 and m_2, respectively. Use conservation of energy, conservation of momentum, and Eq. 14.57, $E^2 - (c_0 p)^2 = (mc_0^2)^2$, to calculate the energy and momentum magnitude of particle 1 in the Earth reference frame in terms of m_{orig}, m_1, m_2, and c_0. ••

66. Which of the following forms of energy contribute to the mass of a gas molecule: (*a*) energy due to the molecule rotating about its center of mass, (*b*) energy due to compression of atoms in the molecule, (*c*) energy due to the molecule's translational motion relative to Earth? ••

67. At the Large Hadron Collider in Switzerland, two high-energy protons collide to create new particles. Prior to collision, each proton is accelerated to an energy of 7000 GeV in the Earth reference frame. (*a*) What is the speed of each proton? (*b*) What is the maximum mass possible for any particle created in the collision? The proton mass is 938 MeV/c_0^2; 1 GeV = 1000 MeV. ••

68. An electron e^- and positron e^+ moving at the same speed in the Earth reference frame collide head-on and produce a proton p and an antiproton $\bar{p}$. The electron and positron have the same mass. The proton and antiproton also have the same mass. The mass of the proton is 1836.15 times the mass of the electron. Calculate, in the Earth reference frame, the minimum value possible for the ratio of the electron's kinetic energy to its internal energy in order to have the reaction $e^- + e^+ \rightarrow p + \bar{p}$ take place. ••

69. Antihydrogen is the only antimatter element that has been produced in the laboratory, albeit just a few atoms at a time. Each antihydrogen atom consists of a positron in orbit around an antiproton and has the same atomic mass as hydrogen. If an antihydrogen atom collides with a hydrogen atom, they annihilate each other and create gamma radiation. (*a*) What minimum amount of energy is released in this process? (*b*) If this energy could be harnessed in a matter-antimatter automobile engine, how far could a car travel on 10 mg each of antihydrogen and hydrogen? At highway speeds, a typical automobile expends about 2.5×10^3 J per meter. •••

Additional Problems

70. Galaxy A moves away from galaxy B at $0.600c_0$ relative to B. A spaceship leaves a planet in galaxy A traveling at $0.500c_0$ relative to galaxy A. If the direction in which the ship travels is the same as the direction in which A is moving away from B, what do observers in B measure for the ship's speed? •

71. A tree that is 32.6 m tall leans 26.0° from the vertical in the Earth reference frame. How tall is the tree, and at what angle does it lean, in the reference frame of a cosmic ray muon moving parallel to the tree at a speed of $0.382c_0$ relative to Earth? •

72. Consider a searchlight on the ground that casts a spot on a cloud 1000 m overhead. If the searchlight is rotated rapidly—say, 30° in 1 μs—how fast does the spot move in the Earth reference frame? Is this a violation of special relativity? ••

73. Two events occur at different locations on Earth and are simultaneous in the reference frame of an observer standing on Earth halfway between the events. Are these events simultaneous in any other inertial reference frame? ••

74. Pilot A is seated in the middle of a ship that has a proper length of 250 m, and pilot B is seated in the middle of an identical ship. Pilot B passes A at a speed that A measures to be $0.580c_0$. How long does it take B's ship to pass pilot A (*a*) according to A and (*b*) according to B? How long does it take A's ship to pass pilot B (*c*) according to A and (*d*) according to B? ••

75. Relative to Earth, spaceship A travels at $0.732c_0$ away from Earth, and spaceship B travels at $0.914c_0$ toward Earth along the same straight line. (*a*) How fast does A move according to an observer aboard B? (*b*) At $t = 0$ the two ships are separated by 4.5×10^{10} m in the Earth reference frame. At what instant t do they pass each other according to an observer at rest on Earth? ••

76. A proton p_1 moving in the Earth reference frame collides with a proton p_2 that is at rest in that reference frame. The collision creates a particle that has mass m and an internal energy 40 times the internal energy of a proton: $mc_0^2 = 40m_p c_0^2$. The velocity of this new particle in the Earth reference frame is not given. What is the ratio of the energy of p_1 to its internal energy in the Earth reference frame? (Hint: First analyze the collision in the zero-momentum reference frame in which p_1 and p_2 are both in motion, moving in opposite directions at the same initial speed.) ••

77. You are consulting for an episode of *Stars Wars, Part XVI*. As this episode opens, starship Orion zooms past a space station at $t = 0$ ship's time, traveling at a constant $0.6000c_0$ relative to the station. When the Orion's clock reads $t = 1000$ s, her sensors detect a hostile battle cruiser. Based on its observed speed, the cruiser is calculated to be 8.00×10^7 km dead ahead and closing in, traveling on a direct line toward the Orion and the space station at a speed of $0.8000c_0$ relative to the Orion. The Orion immediately sends a light signal to the station warning the crew to evacuate. Your boss says to assume that evacuating the station requires 45.0 min measured in the reference frame of the station, and wants to know if a rewrite is required. ••• CR

78. Your boss wants you to construct a spring that, when compressed, experiences a mass increase of 1/10,000 of 1%. He has no idea whether this is possible, but you decide to write a report listing desirable parameters (spring mass, spring constant, load mass), discussing problems you might encounter in attempting to make such a spring, and considering whether or not the conventional spring equations apply. ••• CR

79. The Lorentz Cup spaceship race is won by the pilot who can travel 1.00×10^9 m in the shortest time interval. The rules state that the start and finish lines must be 10^9 m apart in the ships' reference frame, but the time interval is measured by an observer at rest in the Earth reference frame. There are many racing speeds you might try, such as $0.420c_0$ and $0.980c_0$ relative to the Earth reference frame, but trying a few combinations randomly makes you wonder if there is a better way than trial and error to decide on a strategy for this race. ••• CR



15

Periodic Motion

CONCEPTS

15.1 Periodic motion and energy
15.2 Simple harmonic motion
15.3 Fourier's theorem
15.4 Restoring forces in simple harmonic motion

QUANTITATIVE TOOLS

15.5 Energy of a simple harmonic oscillator
15.6 Simple harmonic motion and springs
15.7 Restoring torques
15.8 Damped oscillations

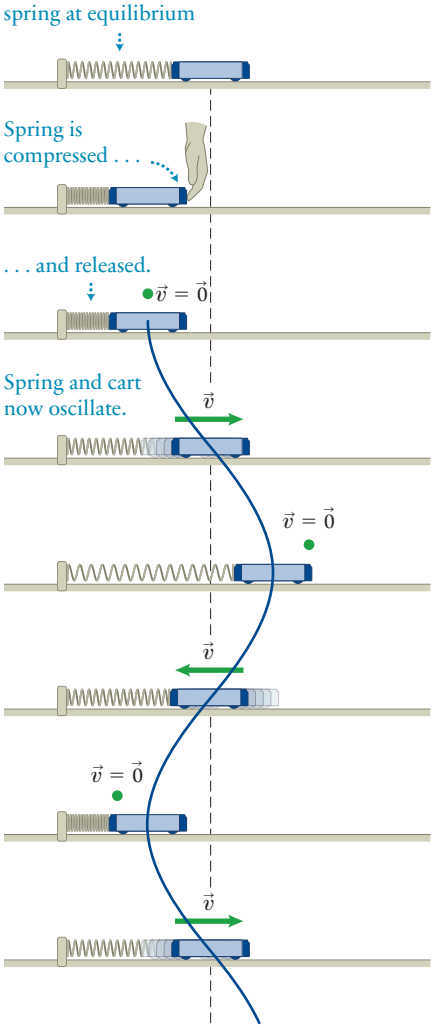
Any motion that repeats itself at regular time intervals is called **periodic motion.** The motion of, say, the Moon revolving around Earth or Earth revolving around the Sun is one type of familiar periodic motion, as is the rotation of the hands of a clock. In this chapter, we are interested in a particular type of periodic motion—namely, back-and-forth periodic motion. Such motion, called either **vibration** or **oscillation** (the two words mean the same thing), is common at all scales in the universe. A rocking chair, a swing, the pendulum of a grandfather clock, the strings on a guitar, the wings of a mosquito all oscillate. At the atomic level, atoms oscillate inside solids. At the cosmic level, the entire universe may oscillate in an ever-repeating cycle of expansion and contraction. The physical natures of these periodic motions are very different from one another, and yet they are all closely related and their mathematical descriptions are always of the same form, differing only in the quantities involved.

In this chapter, we first develop a model for the simplest type of oscillation, called *simple harmonic motion*, and relate it to a periodic motion we are already familiar with: rotational motion. Then we show that any oscillation can be described in terms of simple harmonic motions. After presenting a number of examples, we discuss the effects of energy dissipation during oscillations.

15.1 Periodic motion and energy

Figure 15.1 shows the periodic motion of a spring-cart system. The system, initially at rest, is compressed and then released from rest. After being released, the cart oscillates back and forth along the low-friction track, alternately compressing and stretching the spring. (Ignore friction.)

Figure 15.1 Compressing and releasing the spring cause the cart to oscillate back and forth on the track.



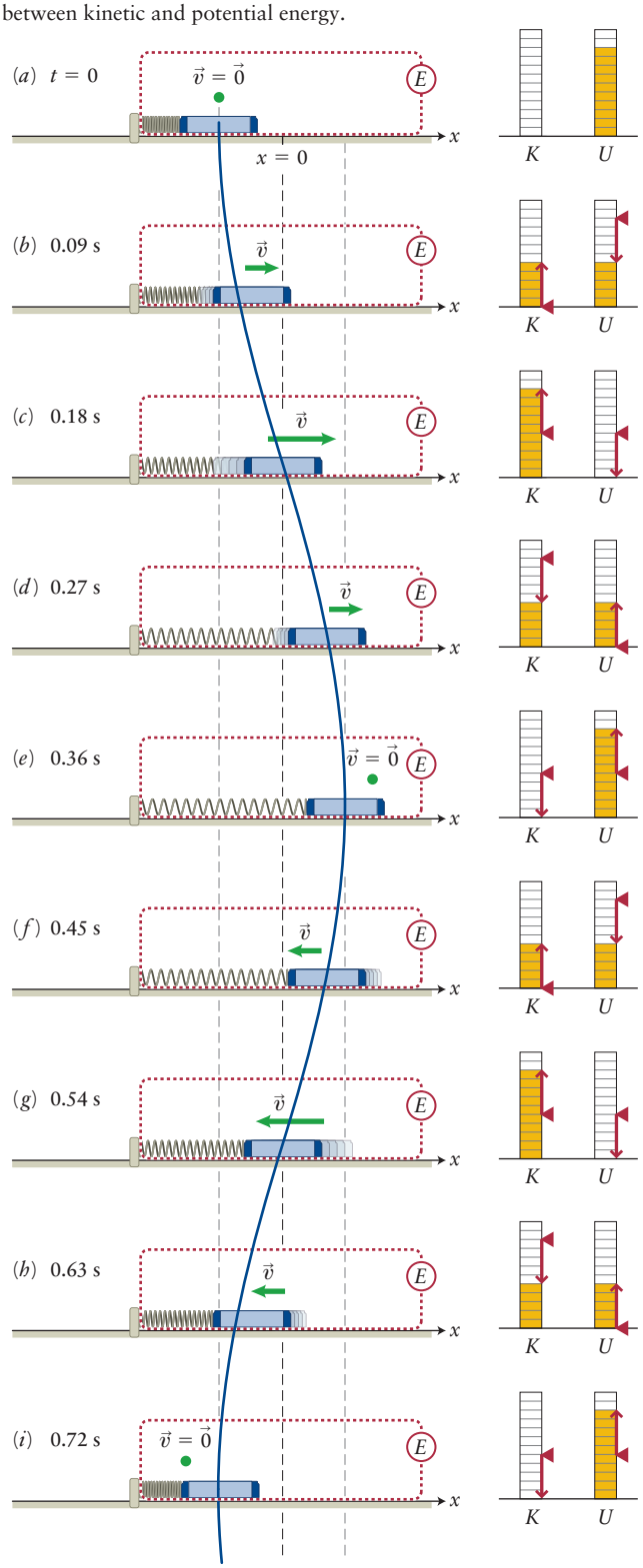
15.1 (*a*) List the forces exerted on the spring-cart system of Figure 15.1 right after it is released, and draw a free-body diagram for each object in the system. (*b*) Which of these forces do work on the system as it oscillates? (*c*) As the hand pushes on the cart and compresses the spring, is the work done by the cart on the spring positive, negative, or zero? (*d*) How does the work done by the cart on the spring compare with the work done by the spring on the cart?

To understand why the motion of the cart in Figure 15.1 is periodic and why the cart doesn't simply return to its initial position and remain there, let us look at the mechanical energy in the cart-spring system. As the spring is compressed by the hand, elastic potential energy is stored in the system. Just before the cart is released, the system is at rest, so the mechanical energy of the system consists entirely of potential energy (**Figure 15.2***a*). After the cart is released, the spring-cart system is closed—none of the forces exerted on it does any work on it—and so the energy of the system must remain constant.

In order to analyze the system's energy changes in more detail, we choose a $+x$ axis that points in the initial direction of motion of the cart and that has its origin at the center of the cart's equilibrium position. As the expanding spring speeds up the cart in the positive x direction, the kinetic energy increases and the potential energy decreases (Figure 15.2*b*). When the cart reaches the equilibrium position at $x = 0$, the potential energy is zero (Figure 15.2*c*). The cart, however, does not stop at this position because its kinetic energy is not zero. The cart therefore overshoots the equilibrium position and begins stretching the spring. As the spring stretches, it exerts on the cart a restoring force that slows the cart down and converts kinetic

CONCEPTS

Figure 15.2 Periodic motion is the result of a continuous exchange between kinetic and potential energy.



energy back to elastic potential energy (Figure 15.2*d*). This continues until the kinetic energy falls to zero, at which point the cart reverses its direction of travel (Figure 15.2*e*). The potential energy is converted back to kinetic energy (Figure 15.2*f*), and the cart again returns to the equilibrium position, this time traveling leftward (Figure 15.2*g*). The potential energy is zero again, but with the cart's kinetic energy being nonzero, it overshoots the equilibrium position once again. Provided no energy is dissipated, the system returns to its initial state (compare Figures 15.2*a* and 15.2*i*). A new cycle begins, and the cart repeats its back-and-forth motion.

15.2 (*a*) In Figure 15.2*e*, the cart's displacement from the equilibrium position is maximum. Is the x component of the cart's acceleration at that instant positive, negative, or zero? (*b*) At which instant(s) in Figure 15.2 is the magnitude of the cart's acceleration greatest? At which instant(s) is it smallest? (Use the blue position curve to answer these questions.)

The time interval it takes to complete a full cycle of the motion is the *period* T (see Section 11.1). For the motion shown in Figure 15.2 the period is equal to the time elapsed between (*a*) and (*i*): $T = 0.72$ s. The inverse of the period, called the **frequency** of the motion, $f = 1/T$, gives the number of cycles completed per second. In Figure 15.2, $f = 1/(0.72\text{ s}) = 1.4\text{ s}^{-1}$ and so the cart completes 1.4 cycles each second.

The magnitude of the cart's maximum displacement from the equilibrium position is called the **amplitude** A of the periodic motion. The amplitude is related to the mechanical energy of the system. In Figure 15.2, for example, the greater the initial compression of the spring, the greater the amplitude and the greater the initial potential energy.

In practice, periodic motion in mechanical systems does not continue indefinitely because some energy is always dissipated (friction in the cart's bearings, air resistance, heating of the spring). This dying out of periodic motion due to dissipation of energy is called *damping*. Initially, we shall ignore damping and, with this simplification, we find that:

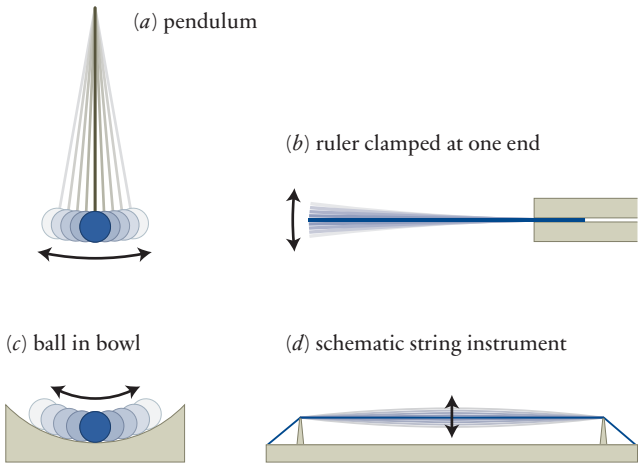
Periodic motion is characterized by a continuous conversion between potential and kinetic energy in a closed system.

All systems that exhibit this behavior have one common feature: a restoring interaction that tends to return the system to equilibrium. Some examples are illustrated in **Figure 15.3**. Each of these systems executes periodic motion when disturbed from its equilibrium position, and in each case the amplitude of the motion is a measure of the mechanical energy of the system.

15.3 For each system in Figure 15.3, identify (*a*) the restoring force and (*b*) the type of potential energy associated with the motion.

CONCEPTS

Figure 15.3 Examples of oscillating systems.



15.2 Simple harmonic motion

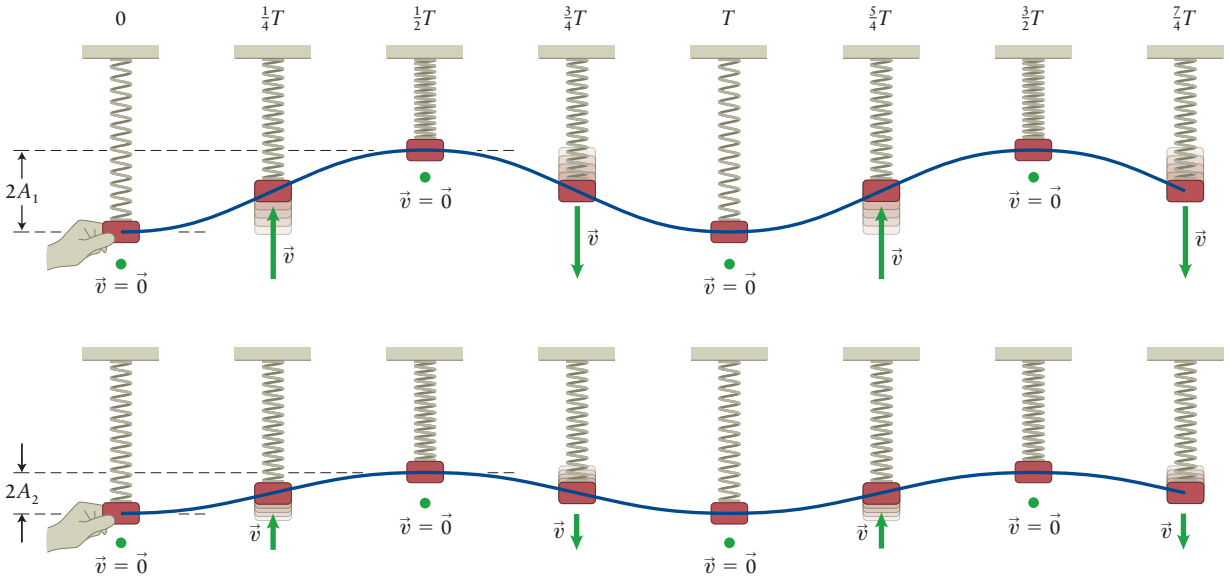
Measuring the period of any oscillating system reveals a surprising fact: When the amplitude of the motion is not too great, the period is independent of the amplitude. For example, an object suspended from a spring has a fixed period of oscillation regardless of its amplitude (**Figure 15.4**). The period of a pendulum, too, is independent of the amplitude of its swing, which is why pendulums are used to regulate the movement of mechanical clocks. Likewise, when you pluck a guitar string, the pitch of the sound, which is determined by the string's period of oscillation, is the same regardless of how far back you pull the string.

A system that exhibits equal periods for all amplitudes is called *isochronous*. Provided their amplitudes are not too great, all the systems shown in Figure 15.3 are isochronous. Individual systems have different periods that are determined by the properties of the system, but any given system has a period that is essentially independent of amplitude.

The graph of displacement versus time for an isochronous oscillating system is shown in **Figure 15.5**. The $x(t)$ curves for an isochronous oscillation have the shape of sine functions. For this reason they are referred to as *sinusoidal curves*. Any periodic motion that yields a sinusoidal $x(t)$ curve is called **simple harmonic motion,** and a system that executes such motion is called a *simple harmonic oscillator*. For small amplitudes, all the systems shown in Figure 15.3 yield this type of $x(t)$ curve, and so all are examples of simple harmonic motion. The motion is called harmonic because in musical instruments harmonic motion is responsible for the sound emitted by the instruments.

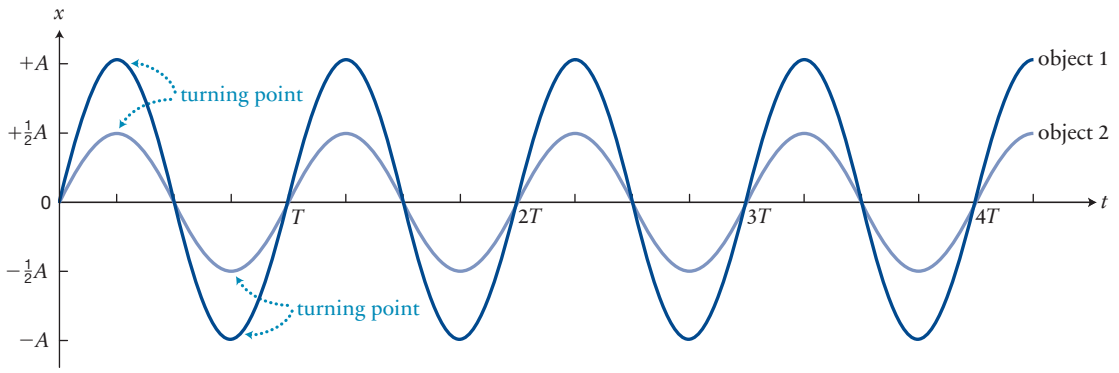
The fact that simple harmonic motion is isochronous allows us to draw some important conclusions. Consider, for example, the two isochronous $x(t)$ curves in Figure 15.5, one with twice the amplitude of the other. These could be the curves of two identical pendulums or any other of the systems shown in Figure 15.3, with one oscillating object moving twice as far out as the other. In one period, object 1 always covers twice the distance that object 2 covers. Because the periods are the same, the velocity of object 1 must always be twice as great as that of object 2 at the same instant in the periodic motion. Neither motion has constant velocity or constant acceleration—the velocities and accelerations change constantly. However, in order for the velocity of object 1 always to be twice

Figure 15.4 The oscillation of an object suspended from a spring is isochronous: The period T does not depend on the amplitude of the oscillation.



CONCEPTS

Figure 15.5 Positions-versus-time graphs for two isochronous objects.



that of object 2, the acceleration of object 1 must, at each instant, be twice as great as the acceleration of object 1.

As you saw in Checkpoint 15.2, the acceleration is greatest when the displacement from the equilibrium position is greatest (that is, at the *turning points*, which are the positions where the object turns around). This tells us that the acceleration of object 1 at its turning point $x_1 = A$ is twice as great as that of object 2 at its turning point $x_2 = \frac{1}{2}A$. Given that the two objects are identical, we conclude that the acceleration of an isochronous object is proportional to the displacement from the equilibrium position. Because the acceleration of an object is proportional to the vector sum of the forces exerted on it, this tells us that the restoring force exerted on the oscillating object is also proportional to the displacement from the equilibrium position. We already know that springs exert linear restoring forces (Hooke's law, Section 8.9). The observation that all systems in Figure 15.3 are isochronous tells us, then, that the restoring forces in all these systems must also be linearly proportional to the displacement from the equilibrium position.

CONCEPTS

An object executing simple harmonic motion is subject to a linear restoring force that tends to return the object to its equilibrium position and is linearly proportional to the object's displacement from its equilibrium position.

15.4 Suppose the spring in Figure 15.1 is compressed twice as much. (*a*) By how much does the mechanical energy of the spring-cart system increase? (*b*) What is the relationship between the amplitude of the oscillation and the mechanical energy in the oscillating system?

Simple harmonic motion is closely related to a periodic motion we have already studied: circular motion at constant speed. To see how the two are related, imagine projecting the shadow of a small sphere in circular motion at constant speed onto a screen perpendicular to the plane of the motion (**Figures 15.6** and **15.7**). While the sphere sweeps out a circular trajectory, its shadow oscillates up and down on the screen. The position of the shadow at any instant is given by the vertical component of the sphere's position vector at that instant. If the sphere begins on the horizontal axis and moves with a constant rotational speed ω, the angle ϕ between the sphere's position vector and the horizontal axis grows steadily with time: $\phi = \omega t$. The position of the sphere's shadow on the screen is then described by $A \sin(\omega t)$, where A is the radius of the sphere's trajectory. Because it is described by a sinusoidal function, the projection of circular motion at constant speed is simple harmonic motion.

The correspondence between circular motion at constant speed and simple harmonic motion can be demonstrated experimentally by projecting the shadows of an object in simple harmonic motion and an object in circular motion at constant speed on a screen, as in Figure 15.7. If the rotational speed is adjusted so that its period equals that of the oscillation, the shadows of the two objects move synchronously.

15.5 What are (*a*) the direction of the velocity of the shadow on the screen in Figure 15.6 and (*b*) the direction of the shadow's acceleration?

Figure 15.6 If we project the shadow of a ball moving in a circle at constant speed onto a screen perpendicular to the plane of the motion, the shadow moves up and down in simple harmonic motion.

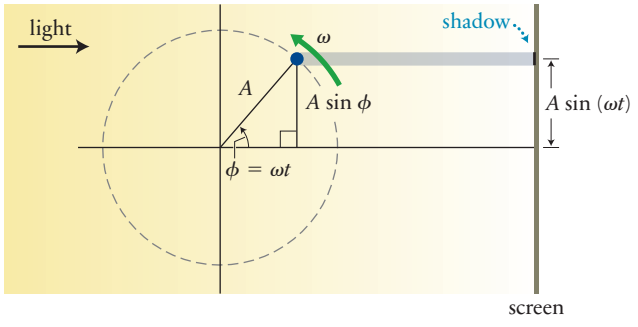
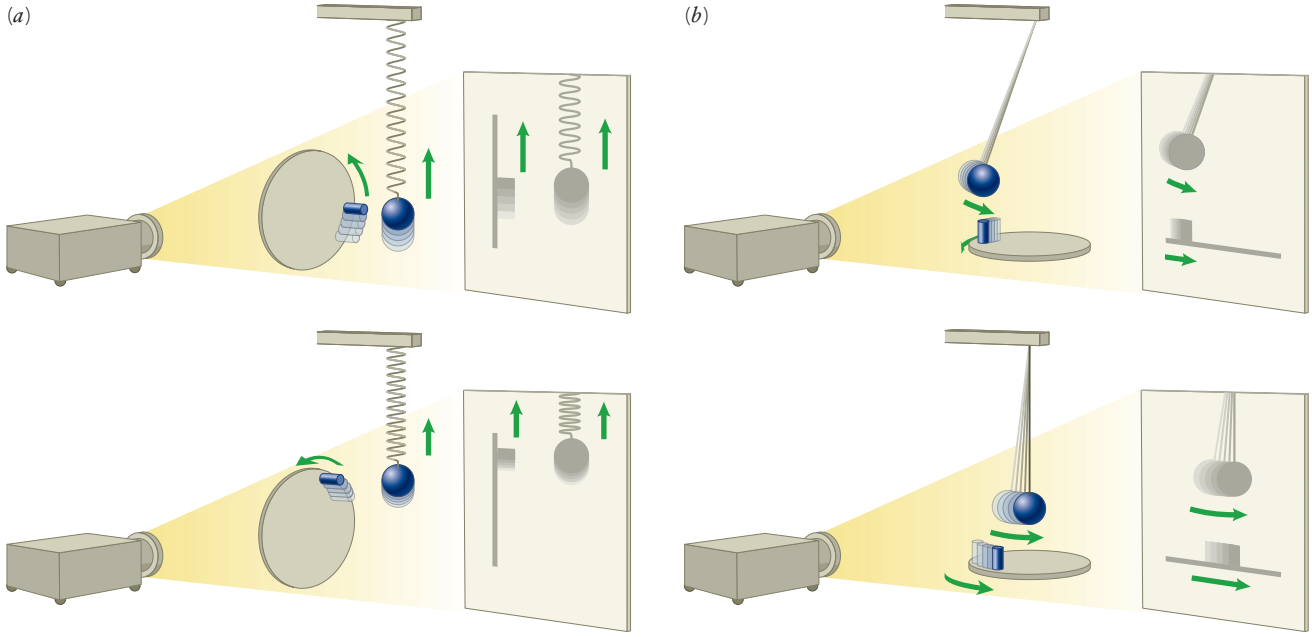


Figure 15.7 Correspondence between circular motion at constant speed and simple harmonic motion for (*a*) an object suspended from a spring and (*b*) a pendulum.



15.3 Fourier's theorem

Simple harmonic motion is important because, as Figure 15.3 suggests, it is very common. Another reason it is important is that, according to a rule known as **Fourier's theorem,*** any periodic function can be written as a sum of sinusoidal functions of frequency $f_n = n/T$, where $n \geq 1$ is an integer and T is the period of the function. This means that *any* periodic motion can be treated as a superposition of simple harmonic motions. So, if we understand simple harmonic motion, we can deal with any other type of periodic motion by breaking it down into its simple harmonic components. Because any motion that is nonrepetitive can be thought of as periodic motion that has an infinitely long period, Fourier's theorem has broad implications: The time-dependent motion of *any* system can be expressed in terms of sinusoidal functions.

Consider, for example, the periodic motion represented by the nonsinusoidal curve in **Figure 15.8*a*** on the next page. Although very different from simple harmonic motion, this motion can be expressed as the sum of a few simple harmonic components, shown in the sinusoidal curves of Figure 15.8*b*. The first component has the same frequency $f_1 = 1/T$ as the original periodic function; this frequency, the lowest in the sum, is called the **fundamental frequency** or **first harmonic.** All other components, called **higher harmonics,** are at integer multiples of the fundamental frequency: $f_n = nf_1 = n/T$.† When these higher harmonics are added to the first harmonic, the shape of the resulting periodic function is not sinusoidal, but the period remains the same (Figure 15.8*c*). By adjusting the amplitude of each harmonic, we can make the sum of harmonics fit the original periodic function.

The breaking down of a function into harmonic components is called *Fourier analysis,* and the resulting sum of sinusoidal functions is called a *Fourier series*. **Figure 15.9** on the next page shows one way of graphically representing the Fourier series: The amplitudes A_n of all the harmonics are plotted against the frequencies of the harmonics. As you saw in Checkpoint 15.4, the energy of a simple harmonic oscillator is proportional to the square of the amplitude. For this reason, it is customary to plot A_n^2 as a function of frequency, as shown in Figure 15.9*c*; such a plot is called the **spectrum** of the periodic function.

The spectrum shows the contribution of each harmonic to the original periodic function. A similar type of visual display of frequency content is found on some stereophonic components. Displays such as the one

CONCEPTS

*After the French mathematician Jean-Baptiste Joseph Fourier (1768–1830).

†In general, an infinite sum of sines and cosines is required to break down arbitrary periodic functions. Only periodic functions for which $f(t) = -f(-t)$ (that is, they are odd in t, such as the one shown in Figure 15.8) can be analyzed with only sines (which are also odd in t).

Figure 15.8 Fourier analysis of a periodic function into sinusoidal functions. The period of each function is indicated by the shading. (*a*) Original periodic function. (*b*) The first through fourth harmonics used to describe the function in (*a*). (*c*) Successive sum of those harmonics, showing an increasingly improving match to the original periodic function.

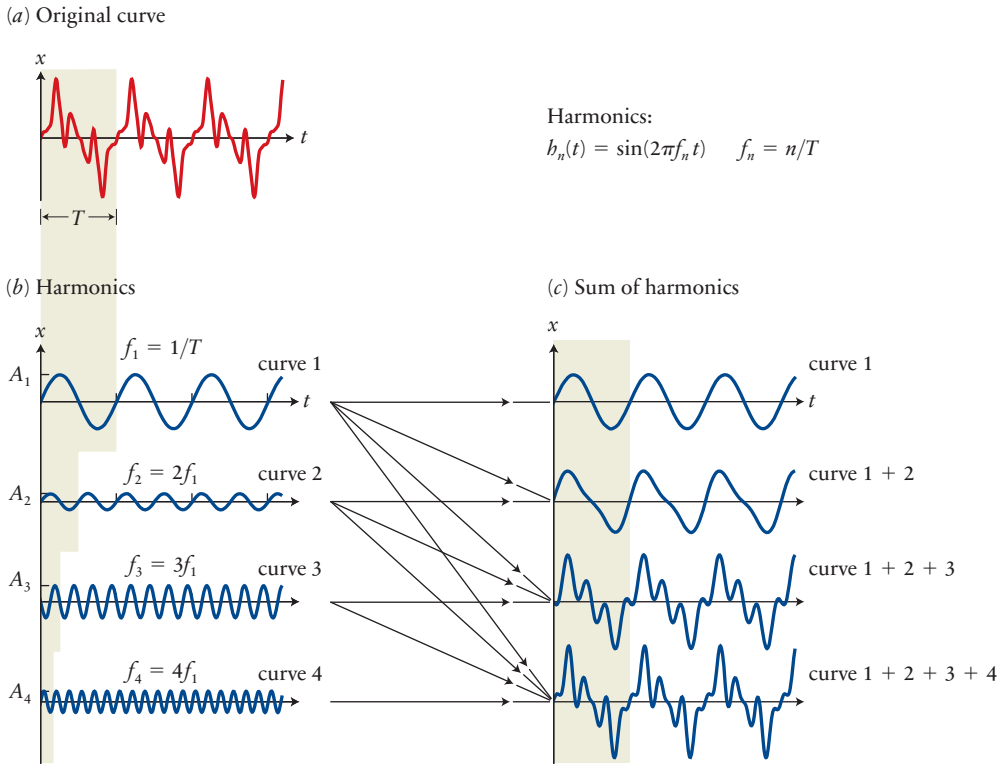
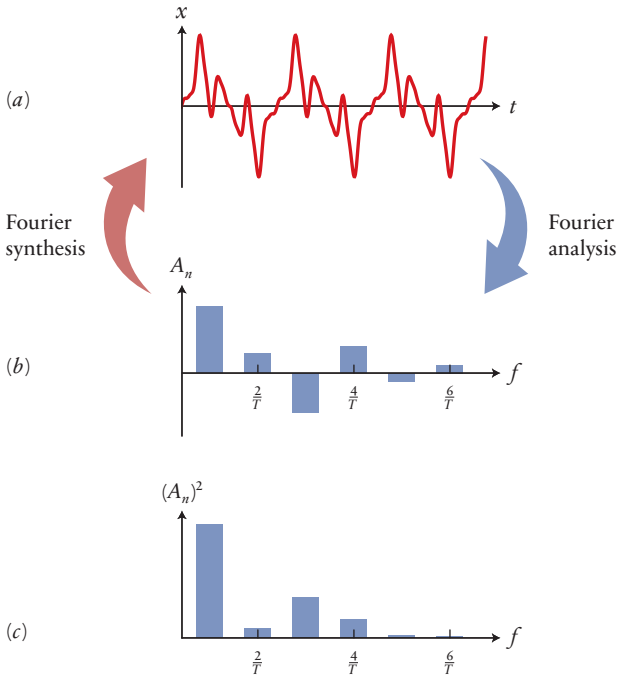


Figure 15.9 (*a*) Periodic function and (*b*) amplitudes of the harmonic functions in the Fourier series of the function shown in part *a*. (*c*) Spectrum of the periodic function, obtained by plotting the squares of the amplitudes.

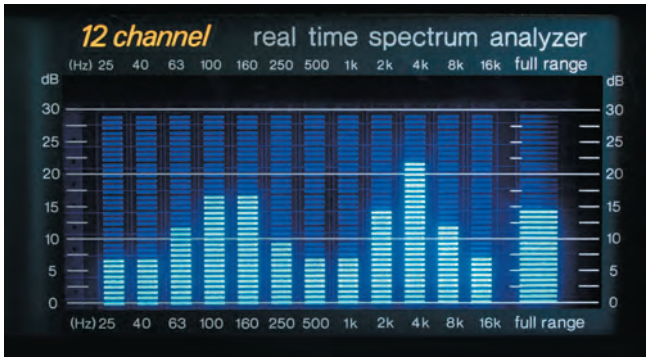


shown in **Figure 15.10** provide a visual representation of the frequency content of the sound being played.

Figure 15.11 shows some important relationships between a periodic function and its spectrum. The period T of the function determines the frequency of the first harmonic. In other words, the duration of the slowest component of the function determines the lowest frequency in the spectrum. Conversely, the duration T_{min} of the fastest component in the function determines the highest frequency in the spectrum.

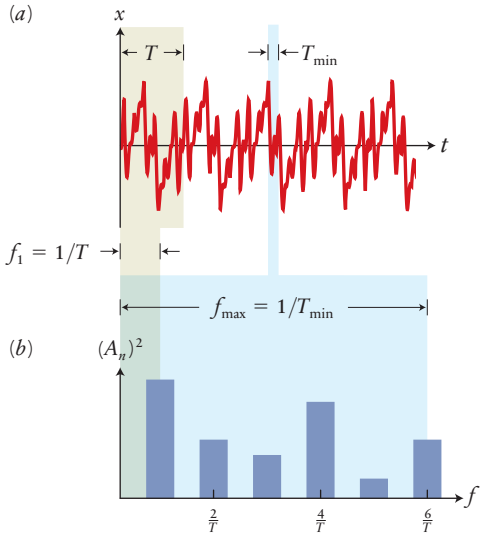
Fourier analysis has many applications, ranging from electronic signal processing to chemical analysis

Figure 15.10 Spectrum analyzer display on a stereo component.



CONCEPTS

Figure 15.11 Relationship between a periodic function and its spectrum.



and voice recognition. The ear functions as a Fourier analyzer because the inner ear breaks down the oscillations produced by sound into simple harmonic components, each producing a nerve impulse proportional to the amplitude of that component.

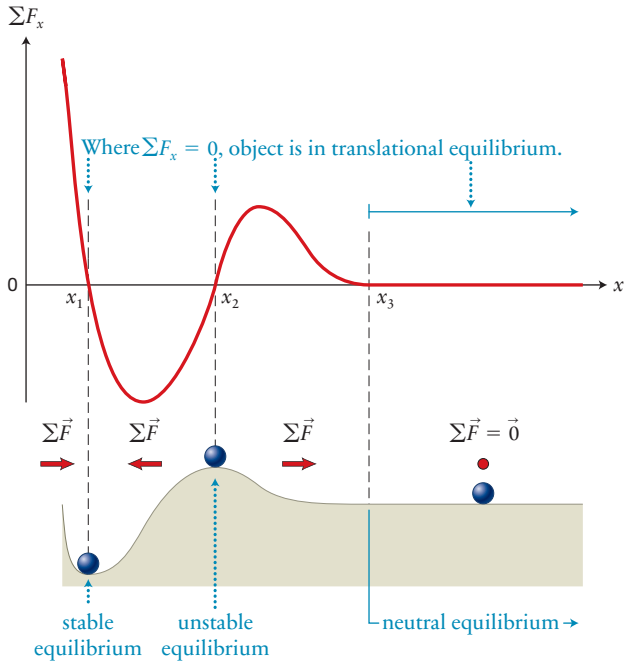
The inverse of Fourier analysis—the creation of periodic functions by adding sinusoidal functions together—is called *Fourier synthesis*. A musician's electronic synthesizer, for example, produces a sum of harmonics that add up to produce different periodic functions. When the amplitudes of the different harmonics are adjusted so that they match the frequency content of real instruments, the synthesizer simulates the sounds of many musical instruments.

15.6 (*a*) What does the spectrum of a single sinusoidal function of period T look like? (*b*) As T is increased, what change occurs in the spectrum?

15.4 Restoring forces in simple harmonic motion

Periodic motion takes place about a position of translational or rotational equilibrium (see Section 12.5) and requires a restoring force that tends to return the object to the equilibrium position. Consider, for example, an object moving along an x axis and subject to the vector sum of forces whose x component is shown graphically in **Figure 15.12**. The object is in translational equilibrium at each position where the vector sum is zero: at x_1, x_2, and all values of $x > x_3$. Can the object oscillate about any of these positions? From the graph we see that ΣF_x is positive to the left of x_1, and so the vector sum $\Sigma\vec{F}$ points toward x_1. To the right of x_1, ΣF_x is negative, and so the vector sum of forces

Figure 15.12 Whenever the vector sum of the forces exerted on an object is zero, the object is in translational equilibrium. This graph shows stable, unstable, and neutral equilibria.



again points toward x_1. The forces exerted on the object tend to return the object to x_1 whenever it moves away from x_1. Around x_1 the vector sum of forces "restores" the translational equilibrium so that the object can oscillate around x_1.

Because the vector sum of the forces always pushes the object back to x_1, this equilibrium position is said to be *stable*. You can visualize this position by imagining a ball at the bottom of a bowl: When the ball is displaced in any direction, a restoring force accelerates it back toward the bottom.

On either side of equilibrium position x_2, the direction of the vector sum of the forces is such that the object accelerates away from the equilibrium position. Such an equilibrium position is called *unstable* because the forces tend to amplify any disturbance from equilibrium. Imagine balancing a ball on top of a hill: The slightest displacement of the ball in any direction is enough to make it accelerate away from the equilibrium position.

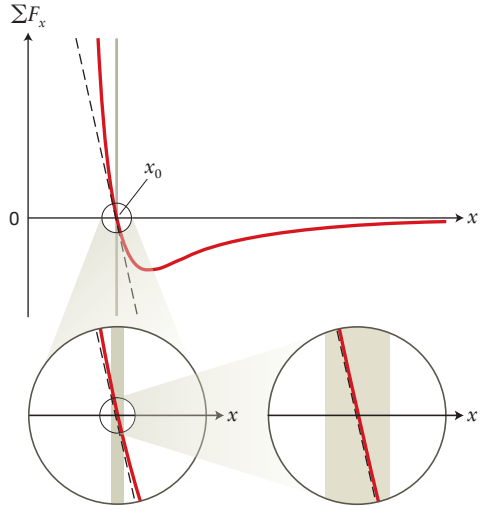
Positions for which $x > x_3$ represent a *neutral* equilibrium: Moving the ball in any direction has no effect on its subsequent acceleration.

What all this tells us is that objects tend to oscillate about *stable* equilibrium positions because the forces about such positions tend to restore the equilibrium. A consequence of the restoring forces about stable equilibrium positions is:

In the absence of friction, a small displacement of a system from a position of stable equilibrium causes the system to oscillate.

CONCEPTS

Figure 15.13 For sufficiently small displacements near a stable equilibrium position, the restoring force is linear in the displacement.



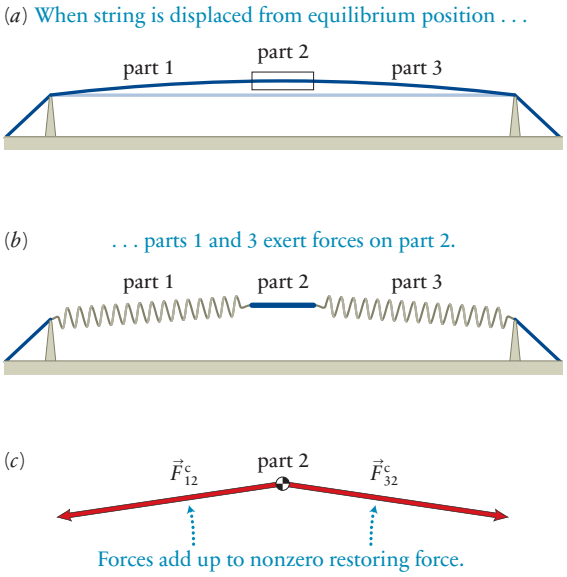
15.7 (*a*) Are there any equilibrium positions in the sum-of-forces-versus-distance graph in **Figure 15.13**? If so, is the equilibrium stable, unstable, or neutral? (*b*) Compare the magnitude of the restoring force for equal displacements on either side of x_0 in Figure 15.13.

Figure 15.13 shows why so many systems move in simple harmonic motion. As we saw in Section 15.2, simple harmonic motion requires *linear* restoring forces—that is, forces that are linearly proportional to the displacement from the equilibrium position. In Checkpoint 15.7*b* you saw that the restoring force about the stable equilibrium position x_0 in Figure 15.13 is not linear. However, as illustrated by the successive enlargements in Figure 15.13, the curve, which gives the relationship between force and position, becomes indistinguishable from a straight line for small displacements from the equilibrium position x_0 (say, within the shaded region). In other words:

For sufficiently small displacements away from the equilibrium position, restoring forces are always linearly proportional to the displacement. Consequently, for small displacements, objects execute simple harmonic motion about any stable equilibrium position.

The restoring forces in the examples in Figure 15.3 fall into two categories: elastic (*b*, *d*) and gravitational (*a*, *c*). We saw in Chapter 8 that, for small displacements from the equilibrium position, the elastic restoring force of a spring is linearly proportional to the displacement. The same is true for other elastic restoring forces. **Figure 15.14**, for example, shows a string displaced from its equilibrium position. Each part of the stretched string is subject to elastic (contact) forces

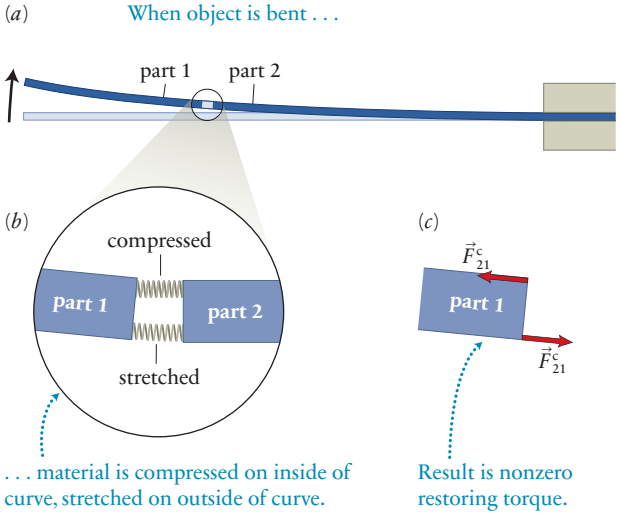
Figure 15.14 Cause of the restoring force for a taut string displaced from equilibrium.



exerted by the parts immediately adjacent to it. Part 2 in the middle of the string in Figure 15.14*b*, for instance, is subject to a force exerted by part 1 and a force exerted by part 3. In the equilibrium position, when the string is straight, the elastic forces exerted by parts 1 and 3 on part 2 add to zero. As the string is displaced, however, the elastic forces exerted by the two neighboring parts no longer line up (Figure 15.14*c*) and their vector sum provides a restoring force that tends to return the string to its equilibrium position. For small displacements, as we just saw, this restoring force is linearly proportional to the displacement from the equilibrium position.

The same line of reasoning applies to the restoring forces that arise when materials are twisted or bent. **Figure 15.15** shows that bending an object compresses

Figure 15.15 Cause of the restoring torque on a beam bent away from its equilibrium shape.

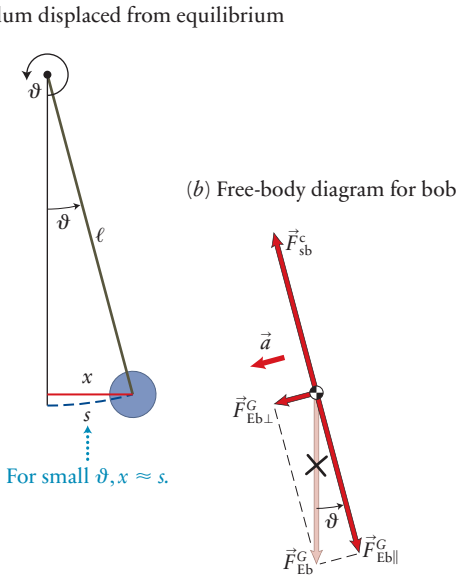


CONCEPTS

the material that lies on the inside of the curve and stretches the material that lies on the outside of the curve. The resulting internal elastic forces cause torques on each part of the object, torques that tend to return the object to its equilibrium shape. For small displacements, the internal elastic forces are linearly proportional to the displacement, and consequently the restoring torque, too, is linearly proportional to the (rotational) displacement.

Let us next examine a pendulum, where the restoring force is due to gravity. **Figure 15.16*a*** shows a pendulum bob of mass m suspended from a string of length ℓ. If the bob is much more massive and much smaller than the string that supports the bob, we can ignore the mass of the string and treat the bob as a particle. Such a pendulum is called a *simple pendulum*. The bob is pulled back over an arc of length s. As the free-body diagram in Figure 15.16*b* shows, the bob is subject to two forces: a contact force exerted by the string and a gravitational force. The restoring force is provided by the component of the gravitational force perpendicular to the string, $\vec{F}^G_{\text{Eb}\perp}$. From Figure 15.16*b*, we see that the magnitude of $\vec{F}^G_{\text{Eb}\perp}$ is equal to $mg\sin\vartheta$ and from Figure 15.16*a* that $\sin\vartheta = x/\ell$, and so the magnitude of the restoring force exerted on the bob is $mg\sin\vartheta = (mg/\ell)x$. Provided the rotational displacement is small, the bob's displacement from its equilibrium position along the circular arc s is approximately equal to the bob's horizontal displacement x, $s \approx x$. For small angles, therefore, the restoring force is linearly proportional to the bob's displacement.

Figure 15.16 For small angles the horizontal displacement x of a pendulum is almost the same as the length of the arc s along which it is displaced.



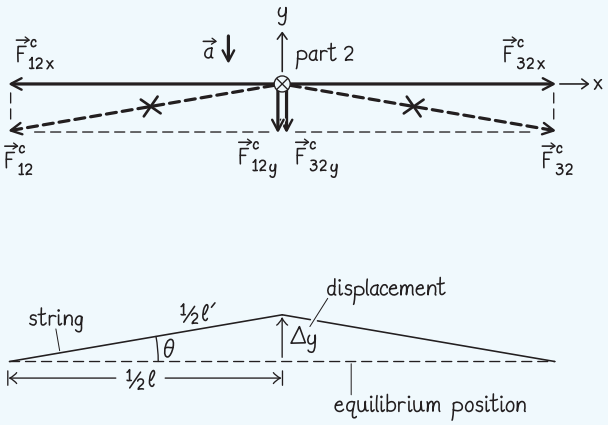
EXAMPLE 15.1 Displaced string

Show that for small displacements the restoring force exerted on part 2 of the displaced string in Figure 15.14 is linearly proportional to the displacement of that part from its equilibrium position.

❶ **GETTING STARTED** Figure 15.14*c* shows the forces exerted by parts 1 and 3 on part 2 when the string is displaced from its equilibrium position. I'll assume that these forces are much greater than the force of gravity exerted on part 2 so that I can ignore gravity in this problem. The force that pulls the string away from the equilibrium position is not shown, which means the string has been released after being pulled away from the equilibrium position.

I begin by making a free-body diagram and choosing a set of axes (**Figure 15.17**). The x components of the forces $\vec{F}^c_{12}$ and $\vec{F}^c_{32}$ cancel; the sum of the y components is the restoring force. The magnitude of the y components is determined by the angle θ between the x axis and either $\vec{F}^c_{12}$ or $\vec{F}^c_{32}$.

Figure 15.17



I also make a sketch of the displaced string, showing the displacement Δy of part 2 from its equilibrium position. I denote the length of the string in its equilibrium position by ℓ and the length of the displaced string by ℓ'.

❷ **DEVISE PLAN** The forces $\vec{F}^c_{12}$ and $\vec{F}^c_{32}$ are equal in magnitude and their y components are determined by $\sin\theta$, which is equal to $\Delta y/\left(\frac{1}{2}\ell'\right)$. If the displacement is small, I can assume that the length of the string doesn't change much from its equilibrium value, so that $\ell \approx \ell'$. Because the forces $\vec{F}^c_{12}$ and $\vec{F}^c_{32}$ are proportional to the tension in the string, I can also consider these forces to be constant. Using this information, I can express the restoring force in terms of the displacement Δy.

❸ **EXECUTE PLAN** From my sketch I see that the restoring force is $F^c_{12y} + F^c_{32y}$. Because $\vec{F}^c_{12}$ and $\vec{F}^c_{32}$ are equal in magnitude, I can write the sum of the y components as $2F^c_{12y} = 2F^c_{12}\sin\theta$. I also know that $\sin\theta = \Delta y/\left(\frac{1}{2}\ell'\right) \approx \Delta y/\left(\frac{1}{2}\ell\right)$. Combining these two relationships, I obtain for the magnitude of the restoring force: $F_{\text{restoring}} = 2F^c_{12y} \approx (4F^c_{12}/\ell)\Delta y$. For small displacements, the term in parentheses is constant and so the restoring force is, indeed, proportional to the displacement Δy. ✔

(Continued)

CONCEPTS

4 EVALUATE RESULT I made two assumptions to derive my answer. The first is that gravity can be ignored. Indeed, taut strings tend to be straight, indicating that gravity (which would make the strings sag) doesn't play an appreciable role. The other assumption I made was that the length of the string doesn't change much when it is displaced from its equilibrium position. This assumption is also justified because the displacement of a string tends to be several orders of magnitude smaller than the string length.

15.8 Using a calculator, determine the percent error in the approximation $\sin\theta \approx \theta$ for polar angles of 1°, 5°, 10°, and 20°. (Note: For this "small-angle" approximation, θ must be expressed in radians.)

In Section 15.1, I showed how oscillations arise from a continuous conversion between kinetic energy and potential energy. Another way to look at oscillations is to say:

Oscillations arise from an interplay between inertia and a restoring force.

When an object that is subject to a restoring force is displaced from its equilibrium position, the restoring force accelerates it back toward the equilibrium position. Once it gets there, however, it has a nonzero velocity, and its inertia causes it to overshoot the equilibrium position.

This picture of an interplay between inertia and restoring force allows us to make some qualitative predictions about the period of an oscillator. The greater the mass of the oscillating object, the harder it is to accelerate and so the longer the period.* Increasing the restoring force has the opposite effect: The greater the restoring force, the greater the object's acceleration and the shorter the period—a stiff spring causes a much more rapid oscillation than a soft, sloppy one.

The period of an oscillating object increases when its mass is increased and decreases when the magnitude of the restoring force is increased.

This relationship does not hold for pendulums, however, because for a pendulum the restoring force and mass are proportional to each other. Increasing the mass of a pendulum bob makes the pendulum harder to accelerate but increases the magnitude of the gravitational force exerted on the pendulum bob—the restoring force—in proportion (see Figure 15.16*b*). The two effects cancel, and so:

The period of a pendulum is independent of the mass of the pendulum.

15.9 (*a*) What effect, if any, does increasing the length ℓ of a simple pendulum have on its period? (*b*) A pendulum and an object suspended from a spring are taken to the Moon, where the acceleration due to gravity is smaller than that on Earth. How does the period of the pendulum on the Moon compare with its period on Earth? (*c*) How does the period of the suspended object on the Moon compare with its period on Earth?

*You can demonstrate this as follows: Position a plastic or thin wooden ruler on a table so that about half its length overhangs the edge of the table. While pressing the ruler down on the table, pull the free end down and release, noting the rapid oscillations. Next attach an eraser or other small object on top of the overhanging part of the ruler and repeat the experiment. Because of the increased mass, the period is now much longer.

CONCEPTS

Self-quiz

1. Is an oscillating object in translational equilibrium?
2. A pendulum bob swings through a circular arc defined by positions A and D in **Figure 15.18**. (*a*) At A, B, and C, what are the possible directions of the velocity of the bob and of the restoring force exerted on it? (*b*) What is the direction of the bob's acceleration at A and C?
3. The cart-spring system shown in **Figure 15.19** is undergoing simple harmonic motion. (*a*) Is the cart's speed greater at A or at B? (*b*) At which of these two positions is the restoring force acting on the cart greater? (*c*) At which of these two positions is the cart's acceleration greater?

Figure 15.18

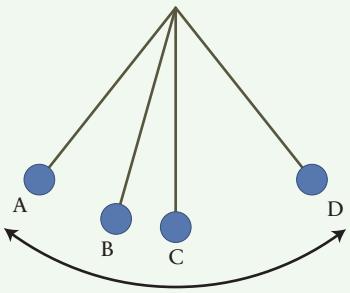


Figure 15.19

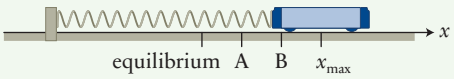
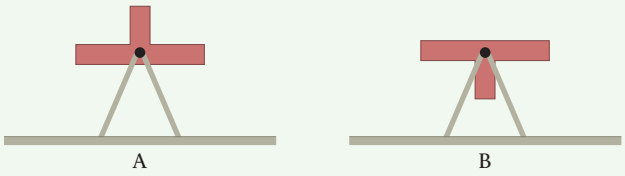


Figure 15.20

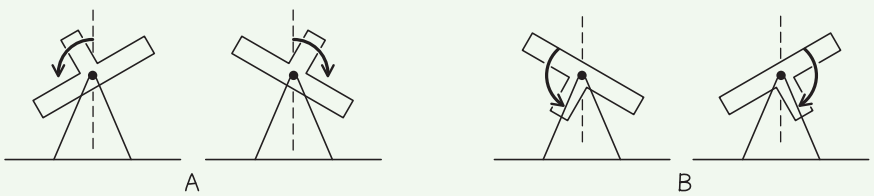


4. A T-shaped wooden structure is balanced on a pivot in the two configurations shown in **Figure 15.20**. Which configuration is in stable equilibrium? Which configuration is likely to oscillate?

Answers

1. No. When an object is in translational equilibrium, the sum of the forces exerted on it is zero. An oscillating object is subject to a variable force that accelerates the object back and forth.
2. (*a*) The velocity is always tangent to the bob's trajectory; the restoring force always points along the arc toward the equilibrium position C. At A, the bob reaches its maximum displacement, so its velocity must be zero; the restoring force points along the arc toward C. At B, the velocity can be in either direction along the arc as the bob swings back and forth, but the restoring force always points along the arc towards C. At C, the velocity can be to the right or left, but the restoring force is zero because this is the equilibrium position. (*b*) At A, the acceleration is tangent along the arc toward C. At C, the bob does have a centripetal radial acceleration directed vertically upward, but no tangential acceleration.
3. (*a*) Once released from the stretched position shown, the cart speeds up as it moves left toward the equilibrium position, so its speed is greater at A than at B. (*b*) The restoring force is proportional to the displacement from the equilibrium position, so it is greater at B than at A. (*c*) The acceleration is proportional to the restoring force, so it, too, is greater at B.
4. If the structure in configuration A is tipped slightly to the left as shown in **Figure 15.21**, the gravitational force exerted on the part to the left of the vertical line through the pivot is greater than that exerted on the part to the right of the pivot, and so the torques caused by these forces make it tip farther in the same direction, away from the equilibrium position. Conversely, if it is tipped to the right, it tips farther to the right. This configuration is unstable, and so the structure does not oscillate. If the structure in configuration B is tipped slightly either way, the torques caused by gravitational forces exerted on the parts on either side of the pivot cause it to move in the opposite direction, back toward the equilibrium position. Configuration B is stable, and so the structure oscillates.

Figure 15.21



15.5 Energy of a simple harmonic oscillator

In Section 15.2, you saw that there is a correspondence between simple harmonic motion and circular motion at constant speed. We begin our quantitative description of simple harmonic motion by exploiting this correspondence.

Figure 15.22 shows a sinusoidally varying, time-dependent function describing some simple harmonic motion represented by a rotating arrow of length A whose tip traces out a circle. The circle is called a **reference circle,** and the arrow is called a **phasor.** The phasor rotates in a counterclockwise direction with a constant rotational speed ω. The vertical component of the phasor varies sinusoidally with time (see Figures 15.6 and 15.7).

Let us first establish a relationship between the rotational speed of the phasor and the frequency f of the corresponding simple harmonic motion. If the phasor completes one revolution in a period T, its rotational speed is

$$\omega \equiv \frac{\Delta\vartheta}{\Delta t} = \frac{2\pi}{T}, \tag{15.1}$$

and the frequency of the corresponding simple harmonic motion is, as noted in Section 15.1,

$$f \equiv \frac{1}{T}. \tag{15.2}$$

The derived SI unit of frequency is the **hertz** (Hz) after the 19th-century German physicist Heinrich Hertz, who produced the first radio waves:

$$1\ \text{Hz} \equiv 1\ \text{s}^{-1}. \tag{15.3}$$

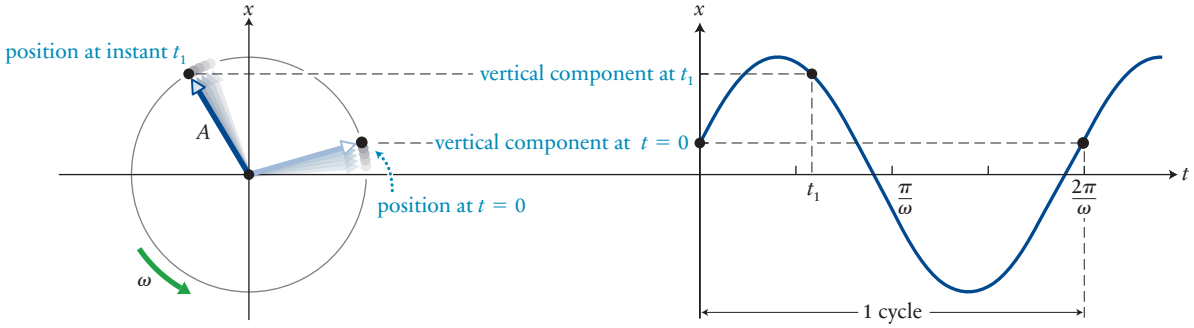
Substituting $1/T$ from Eq. 15.2 into Eq. 15.1, we find

$$\omega = 2\pi f. \tag{15.4}$$

So the rotational speed ω of the phasor is greater than the frequency f of the simple harmonic motion by a factor of 2π. The reason for this factor is that ω measures the change in rotational position per unit time, and in one revolution the rotational position ϑ changes by 2π. Frequency, on the other hand, measures the number of cycles per unit time, and one revolution on the reference circle corresponds to one oscillation cycle (Figure 15.22).

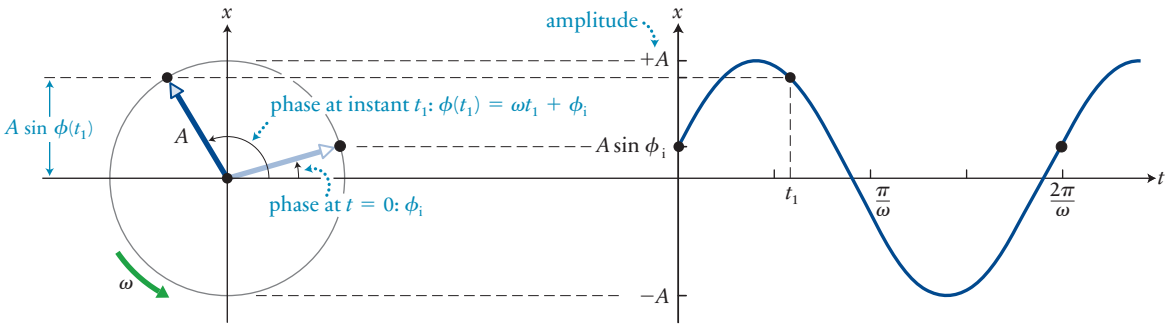
In the context of oscillations, ω is often called the **angular frequency** because it has the same unit (s^{-1} or inverse second) as the frequency f. Keep in mind, however, that the frequency f and the angular frequency ω are not equal (see Eq. 15.4). To facilitate remembering the distinction between the two, we use the SI unit hertz for f and the SI unit s^{-1} for ω.

Figure 15.22 Reference circle and $x(t)$ curve for a simple harmonic oscillator.



QUANTITATIVE TOOLS

Figure 15.23 The phase of a simple harmonic oscillator at any instant is given by the angle between the phasor and the positive horizontal axis measured in the counterclockwise direction.



The rotational position of the tip of the phasor relative to the positive horizontal axis is called the **phase** of the motion (**Figure 15.23**). The phase is represented by the symbol $\phi(t)$. For constant angular frequency, the phase increases linearly with time:

$$\phi(t) = \omega t + \phi_i \quad \text{(constant angular frequency)}, \qquad (15.5)$$

where the constant ϕ_i is the *initial phase* at $t = 0$. The phase $\phi(t)$ tells us at what point in the cycle the motion is; the initial phase ϕ_i is the phase at $t = 0$. Because rotational position is a unitless number, the phase is unitless too. You determine the phase by expressing the angle between the phasor and the positive horizontal axis in radians and dividing by 1 rad.

The vertical component of A in Figure 15.23 can now be written in the form

$$x(t) = A\sin\phi(t) = A\sin(\omega t + \phi_i) \quad \text{(simple harmonic motion)}, \quad (15.6)$$

where the **amplitude** A of the oscillation is equal to the length of the phasor. Equation 15.6 gives a general expression for the position of a simple harmonic oscillator as a function of time when the equilibrium position is at $x = 0$—a pendulum bob, an object suspended from a spring, a taut string. Because the motion is periodic, $x(t)$ has the same value each time the phase increases by 2π. For a crest (where $x(t) = A$ and the sine function in Eq. 15.6 has the value 1), the phase could be $\frac{\pi}{2}$, or $\frac{\pi}{2}$ plus an integer multiple of 2π; for a trough (where $x(t) = -A$ and the sine function in Eq. 15.6 has the value -1), the phase could be $\frac{3\pi}{2}$, or $\frac{3\pi}{2}$ plus an integer multiple of 2π.

EXERCISE 15.2 Oscillation phases

Consider the spring-cart system of Figure 15.1 again. The cart is pulled away from its equilibrium position in the positive x direction and then released at $t = 0$. (*a*) What is the initial phase of the cart's oscillation? What is the phase of the cart's oscillation (*b*) half a period later and (*c*) one period later? What are the cart's initial phases in the following two situations? (*d*) The cart is at its equilibrium position when, at $t = 0$, it is given a sharp blow in the negative x direction so that it starts oscillating. (*e*) The cart is first pulled back a distance d in the positive x direction, and then at $t = 0$ it is given a sharp blow, also in the positive x direction, so that it carries out an oscillation of amplitude $A = 2d$.

SOLUTION (*a*) The cart starts at maximum amplitude in the positive x direction, so the phasor is pointing straight up and the initial phase of the cart's oscillation is $\phi_i = (90°)(2\pi \text{ rad}/360°)/(1 \text{ rad}) = \pi/2$. ✔

(Continued)

QUANTITATIVE TOOLS

(*b*) In one period, the phasor completes a full cycle, so half a period later the phasor has traversed half a cycle (180°). The angle between the phasor and the horizontal axis is thus 90° + 180° = 270°, and thus $\phi(t) = (270°)(2\pi \text{ rad}/360°)/(1 \text{ rad}) = 3\pi/2$. ✔

(*c*) In one period the phasor has completed a full cycle, so the phase is equal to the initial phase $\phi_i + 2\pi = \frac{\pi}{2} + 2\pi = \frac{5\pi}{2}$. ✔

(*d*) If the cart is in the equilibrium position, the phasor must be parallel to the horizontal axis and so the phasor is at 0° or 180°. When the phasor is at 0°, the cart moves in the positive x direction; when it is at 180°, the cart moves in the negative x direction, so the initial phase is $\phi_i = (180°)(2\pi \text{ rad}/360°)/(1 \text{ rad}) = \pi$. ✔

(*e*) The displacement of the cart is given by Eq. 15.6. At $t = 0$: $x(0) = A \sin \phi_i$. I am given that $x(0) = +d$ and $A = 2d$, so $+d = 2d \sin \phi_i$, or $\sin \phi_i = \frac{1}{2}$, and so $\phi_i = \pi/6$ or $5\pi/6$. Because the cart moves in the positive x direction at $t = 0$, the initial phase must be $\phi_i = \pi/6$. ✔

By differentiating Eq. 15.6 with respect to time, we can obtain general expressions for the x components of the velocity and acceleration of a harmonic oscillator:

$$v_x \equiv \frac{dx}{dt} = \omega A \cos(\omega t + \phi_i) \quad \text{(simple harmonic motion)} \tag{15.7}$$

$$a_x \equiv \frac{d^2x}{dt^2} = -\omega^2 A \sin(\omega t + \phi_i) \quad \text{(simple harmonic motion).} \tag{15.8}$$

Equations 15.6–15.8 are the basic kinematic equations for simple harmonic motion. We derived them without considering any particular system—everything followed from the fact that simple harmonic motion corresponds to the projection onto a screen of some object moving in circular motion at constant speed.



15.10 (*a*) What are the algebraic signs of x, v_x, and a_x when the phase $\phi(t) = \omega t + \phi_i$ is between 0 and $\frac{\pi}{2}$? (*b*) Repeat for $\pi < \phi(t) < \frac{3\pi}{2}$.

Substituting Eq. 15.6 into Eq. 15.8, we find that the x component of the acceleration of an object in simple harmonic motion is proportional to the displacement and points in the direction opposite the displacement (that is, toward the equilibrium position):

$$a_x = -\omega^2 x \quad \text{(simple harmonic motion).} \tag{15.9}$$

Substituting the definition of the acceleration into Eq. 15.9, we obtain a second-order differential equation in the variable x:

$$\frac{d^2x}{dt^2} = -\omega^2 x \quad \text{(simple harmonic motion).} \tag{15.10}$$

This equation is called the **simple harmonic oscillator equation** because any system that satisfies this equation undergoes simple harmonic motion. As we shall see in the next two sections, the symbols may take on different meanings for different types of simple harmonic motion, but the mathematical form is always that given in Eq. 15.10.

Multiplying both sides of Eq. 15.9 by the mass of the oscillating object, we have

$$ma_x = -m\omega^2 x \quad \text{(simple harmonic motion).} \tag{15.11}$$

The left side of Eq. 15.11 is equal to the x component of the vector sum of the forces acting on the oscillating object, $ma_x = \Sigma F_x$, and so

$$\Sigma F_x = -m\omega^2 x \quad \text{(simple harmonic motion)}. \tag{15.12}$$

Because m and ω are positive constants, Eq. 15.12 tells us that an object in simple harmonic motion is subject to a force that always points in a direction opposite the object's displacement x from the equilibrium position and is linearly proportional to x—in other words, a *linear restoring force* as we already concluded in Section 15.4.

Let us next examine the mechanical energy of a simple harmonic oscillator, using kinematic Equations 15.6–15.8. The work done by the forces exerted on the oscillating object when it is moving from the equilibrium position to position x is found by integrating the x component of the position-dependent force exerted on the oscillator (Eq. 9.22):

$$W = \int_{x_0}^{x} \Sigma F_x(x)\,dx = -\int_{x_0}^{x} m\omega^2 x\,dx, \tag{15.13}$$

where we have substituted the restoring force given in Eq. 15.12. This work causes the kinetic energy of the oscillating object to change, and so Eq. 15.13 gives us the change in kinetic energy ΔK. Working out the integration thus yields

$$\Delta K = -m\omega^2 \int_{x_0}^{x} x dx = -m\omega^2 [\tfrac{1}{2}x^2]_{x_0}^{x} = \tfrac{1}{2}m\omega^2 x_0^2 - \tfrac{1}{2}m\omega^2 x^2. \tag{15.14}$$

Because the interaction that causes the oscillation is reversible, this change in kinetic energy must be provided by some form of potential energy. (In the examples we discussed, the energy was provided by gravitational potential energy or the elastic potential energy of a spring.) If we include the object that causes the oscillation in our system, the system must be closed, so $\Delta E = \Delta K + \Delta U = 0$. Therefore the change in potential energy of the (closed) system is $\Delta U = -\Delta K$, or, substituting Eq. 15.14,

$$\Delta U = U(x) - U(x_0) = \tfrac{1}{2}m\omega^2 x^2 - \tfrac{1}{2}m\omega^2 x_0^2. \tag{15.15}$$

If we let the potential energy at the equilibrium position be zero, $U(x_0) = 0$, then we see that the potential energy of the oscillating system is $U(x) = \frac{1}{2}m\omega^2 x^2$, and so the mechanical energy is

$$E = K + U = \tfrac{1}{2}mv^2 + \tfrac{1}{2}m\omega^2 x^2 \tag{15.16}$$

or, substituting for x from Eq. 15.6 and for v_x from Eq. 15.7,

$$\begin{aligned} E &= \tfrac{1}{2}m\omega^2 A^2 \cos^2(\omega t + \phi_i) + \tfrac{1}{2}m\omega^2 A^2 \sin^2(\omega t + \phi_i) \\ &= \tfrac{1}{2}\,m\omega^2 A^2 \quad \text{(simple harmonic motion)}, \end{aligned} \tag{15.17}$$

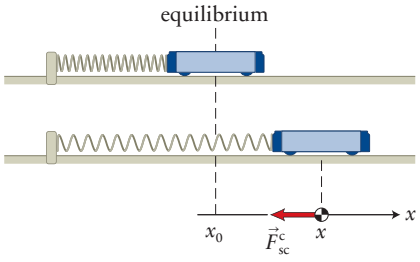
where we have used the fact that $\cos^2\phi + \sin^2\phi = 1$ for all ϕ. Equation 15.17 shows that the mechanical energy of a simple harmonic oscillator is constant and proportional to the square of the amplitude, as we concluded earlier.



15.11 (*a*) Use Eq. 15.15 to determine the maximum potential energy of a simple harmonic oscillator. (*b*) Does your answer to part *a* agree with Eq. 15.17?

15.6 Simple harmonic motion and springs

Figure 15.24 Restoring force exerted by spring on cart when cart is displaced from its equilibrium position x_0.



Whenever an object undergoes reversible deformation as a result of stretching, compressing, bending, or twisting, it exerts an elastic restoring force. Consider, for example, the spring-cart system in **Figure 15.24**. The spring exerts on the oscillating cart a force $F^{c}_{scx} = -k(x - x_0)$ (Eq. 8.20), where k is the spring constant and $x - x_0$ is the displacement of the cart from its equilibrium position at x_0, where the spring is relaxed.

We can use this information about the restoring force to obtain an equation for the cart's motion in terms of the amplitude, angular frequency, and phase of the motion. If we let the origin be at the equilibrium position, $x_0 = 0$ and so

$$F^{c}_{scx} = -kx. \tag{15.18}$$

The equation of motion for the cart in the horizontal direction is then

$$\Sigma F_{cx} = ma_x = -kx, \tag{15.19}$$

where m is the cart's mass, or

$$ma_x = m\frac{d^2x}{dt^2} = -kx \tag{15.20}$$

or

$$\frac{d^2x}{dt^2} = -\frac{k}{m}x. \tag{15.21}$$

The solution to this simple harmonic oscillator equation is a sinusoidal function $x(t)$, and comparison with Eq. 15.10 tells us that $k/m = \omega^2$. Therefore the angular frequency of the oscillation is

$$\omega = +\sqrt{\frac{k}{m}}. \tag{15.22}$$

The plus sign in front of the square root means that we must take the positive square root because ω is always positive.

The motion of the cart is given by Eq. 15.6:

$$x(t) = A\sin\left(\sqrt{\frac{k}{m}}t + \phi_i\right). \tag{15.23}$$

Figure 15.25 Four of the infinite number of solutions that satisfy Eq. 15.21.

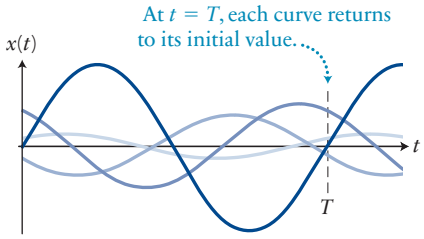


Figure 15.25 shows a number of different sinusoidal functions of the form given by Eq. 15.23 that satisfy Eq. 15.21. All of the functions $x(t)$ shown in Figure 15.25 complete one cycle in a period T, and so they all have the same angular frequency ω, whose value is fixed by properties of the system—the spring constant k and the cart's mass m (Eq. 15.22). However, the peak heights tell you that the curves do not all have the same amplitude A, and the fact that each curve reaches its maximum at a different instant tells you that they have different initial phases ϕ_i. Without information in addition to the values of k and m, we cannot determine the amplitude A and initial phase ϕ_i.

In order to determine A and ϕ_i, we have to know the position and x component of the velocity at some instant. The next examples illustrate how to determine these quantities for some specific situations.

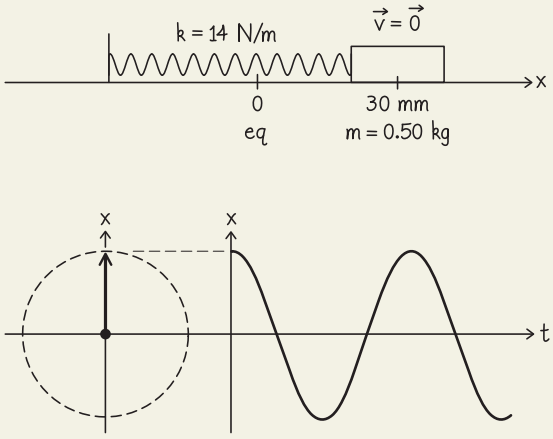
QUANTITATIVE TOOLS

EXAMPLE 15.3 Cart pulled away and released from rest

A cart of mass $m = 0.50$ kg fastened to a spring of spring constant $k = 14$ N/m is pulled 30 mm away from its equilibrium position and then released with zero initial velocity. What are the cart's position and the x component of velocity 2.0 s after being released?

1 GETTING STARTED I begin by making a sketch of the initial condition and representing the oscillation by a reference circle and a graph of position versus time (**Figure 15.26**). I choose the positive x axis to be in the same direction as the cart's initial displacement, with the origin of the axis at the equilibrium position. Because the cart is released with zero initial velocity, its initial kinetic energy at $x = +30$ mm is $K = 0$, and so each time it returns to that position it again has $K = 0$. This tells me that the cart's initial displacement is also its maximum displacement, and so $A = 30$ mm. Because the cart begins at its maximum positive displacement, I know that the oscillation begins with the phasor straight up, and so the initial phase is $\phi_i = +\pi/2$.

Figure 15.26



2 DEVISE PLAN To determine the position of the cart at $t = 2.0$ s, I can use Eq. 15.6. This equation contains the angular frequency ω, which I can determine from Eq. 15.22, and the amplitude A and initial phase ϕ_i, which I already determined. To calculate the x component of the cart's velocity, I can use Eq. 15.7.

3 EXECUTE PLAN From Eq. 15.22, I get for the angular frequency of the oscillation:

$$\omega = \sqrt{\frac{k}{m}} = \sqrt{\frac{14\ \text{N/m}}{0.50\ \text{kg}}} = \sqrt{28\ \text{s}^{-2}} = 5.3\ \text{s}^{-1}.$$

Substituting this angular frequency, $t = 2.0$ s, $A = 30$ mm, and $\phi_i = +\pi/2$ into Eq. 15.6, I obtain for the cart's position at $t = 2.0$ s:

$$x(2.0\ \text{s}) = (30\ \text{mm})\sin\left[(5.3\ \text{s}^{-1})(2.0\ \text{s}) + \frac{\pi}{2}\right]$$

$$= -12\ \text{mm}. \checkmark$$

Substituting the same values into Eq. 15.7, I obtain for the x component of the cart's velocity:

$$v_x(2.0\ \text{s}) = (5.3\ \text{s}^{-1})(30\ \text{mm})\cos\left[(5.3\ \text{s}^{-1})(2.0\ \text{s}) + \frac{\pi}{2}\right]$$

$$= +1.5 \times 10^2\ \text{mm/s}. \checkmark$$

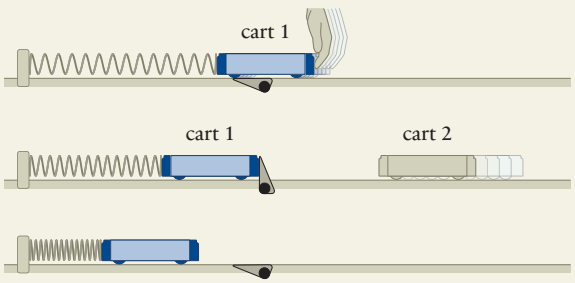
4 EVALUATE RESULT The magnitude of the value I obtained for x is smaller than the amplitude, as it should be. The phase at $t = 2.0$ s is $\phi(t) = (5.3\ \text{s}^{-1})(2.0\ \text{s}) + \pi/2 = 12$, which corresponds to $12/(2\pi) = 1.9$ cycles, so the phasor is in the fourth quadrant, which confirms that x should be negative. The cart moves a distance of $4A = 0.12$ m in one cycle, which takes $2\pi/\omega = 1.2$ s, and so the cart's average speed is $(0.12\ \text{m})/(1.2\ \text{s}) = 0.10$ m/s, indicating that the magnitude of the velocity component I obtain is reasonable. Because the cart is moving in the positive x direction when the phasor is in the fourth quadrant, the x component of the velocity should be positive, as I found.

15.12 (*a*) In Example 15.3 is the initial condition that $v_x = 0$ satisfied by Eq. 15.7? (*b*) Is it correct to say that the x component of the velocity given by Eq. 15.7 must be positive whenever x is negative so as to make sure that the cart moves back to its equilibrium position?

EXAMPLE 15.4 Cart struck with spring already compressed

Cart 1 of mass $m = 0.50$ kg fastened to a spring of spring constant $k = 14$ N/m is pushed 15 mm in from its equilibrium position and held in place by a ratchet (**Figure 15.27**). An identical cart 2 is launched at a speed of 0.10 m/s toward cart 1. The carts collide elastically, releasing the ratchet and setting cart 1 in motion. After the collision, cart 2 is immediately removed from the track. (*a*) What is the maximum compression of the spring? (*b*) How many seconds elapse between the instant the carts collide and the instant the spring reaches maximum compression?

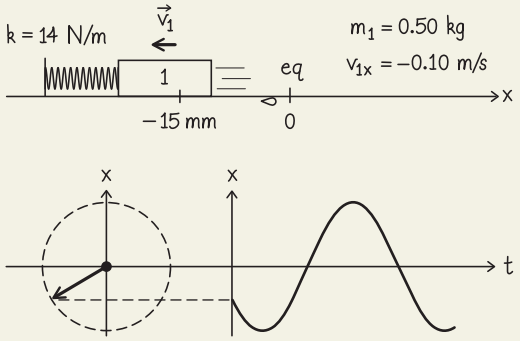
Figure 15.27 Example 15.4.



(Continued)

1 GETTING STARTED If I ignore the effect of the spring during the collision, I can say that the elastic collision interchanges the velocities of the two carts. I make a sketch of the initial condition of cart 1 just before the collision, choosing the positive x axis to the right, the equilibrium position at $x = 0$, and the initial displacement of the cart 15 mm to the left of $x = 0$. I also draw a reference circle and sketch the oscillation resulting from the collision (**Figure 15.28**). With this choice of axis, the x component of the initial velocity of cart 1 is $v_{x,i} = -0.10$ m/s. Because m and k are the same as in Example 15.3, the angular frequency is again $\omega = 5.3\ \text{s}^{-1}$.

Figure 15.28



In contrast to the situation in Example 15.3, the initial displacement of cart 1 is not equal to the amplitude of the oscillation because the collision increases the cart's displacement from the equilibrium position. In other words, cart 1 moves leftward immediately after the collision. It continues moving in this direction until the elastic restoring force building up in the compressing spring causes the cart to stop. This maximum-compression position gives the amplitude of the oscillation.

2 DEVISE PLAN I can determine the value of the amplitude from the mechanical energy of the cart-spring system, which is the sum of the initial potential energy in the spring and the initial kinetic energy of the cart. The potential energy in the spring is given by Eq. 9.23; because the equilibrium position x_0 is at the origin, this equation reduces to $U_{\text{spring}} = \frac{1}{2}kx^2$. The kinetic energy is given by $K = \frac{1}{2}mv^2$. At the position of maximum compression, all of the mechanical energy is stored in the spring, $x = -A$, and so $E_{\text{mech}} = U_{\text{spring}} = \frac{1}{2}kA^2$.

Once I know A, I can determine the initial phase from Eq. 15.6. I can then use that same equation to solve for t at the position of maximum compression, where $x = -A$.

3 EXECUTE PLAN (*a*) The initial kinetic and potential energies of the cart-spring system are

$$K = \tfrac{1}{2}(0.50\ \text{kg})(-0.10\ \text{m/s})^2 = 0.0025\ \text{J}$$

$$U = \tfrac{1}{2}kx^2 = \tfrac{1}{2}(14\ \text{N/m})(-15\ \text{mm})^2\left(\frac{1.0\ \text{m}}{1000\ \text{mm}}\right)^2 = 0.0016\ \text{J}$$

so

$$E_{\text{mech}} = K + U = (0.0025\ \text{J}) + (0.0016\ \text{J}) = 0.0041\ \text{J}.$$

At the position of maximum compression, all of this energy is stored in the spring, and so $\frac{1}{2}kA^2 = 0.0041$ J, or with the k value given,

$$A = \sqrt{\frac{2(0.0041\ \text{J})}{14\ \text{N/m}}} = 0.024\ \text{m} = 24\ \text{mm}. \checkmark$$

(*b*) Substituting the value for A determined in part *a* and the initial condition $x_i = -15$ mm at $t = 0$ into Eq. 15.6, I obtain

$$x(0) = A\sin(0 + \phi_i) = (24\ \text{mm})\sin\phi_i = -15\ \text{mm}$$

$$\sin\phi_i = \frac{-15\ \text{mm}}{24\ \text{mm}} = -0.63 \quad \text{or} \quad \phi_i = \sin^{-1}(-0.63).$$

Two initial phases satisfy this relationship, $\phi_i = -0.68$ and $\phi_i = -\pi + 0.68 = -2.5$, but only the latter gives a negative x component of the velocity (see Eq. 15.7), as required by the initial condition.

At the first instant of maximum compression, $\sin(\omega t + \phi_i) = -1$, which means $\omega t + \phi_i = -\frac{1}{2}\pi$. Solving for t yields $t = \left(-\frac{1}{2}\pi - \phi_i\right)/\omega = \left[-\frac{1}{2}\pi - (-2.5)\right]/(5.3\ \text{s}^{-1}) = 0.17$ s. ✔

4 EVALUATE RESULT At 24 mm, the amplitude is greater than the 15-mm initial displacement from the equilibrium position, as I would expect. From the reference circle part of my sketch I see that it takes about one-eighth of a cycle to go from the position of impact to the position of maximum compression. From Eq. 15.1 I see that the number of seconds needed to complete one cycle is $T = 2\pi/\omega = 2\pi/(5.3\ \text{s}^{-1}) = 1.2$ s, and so the 0.17-s value I obtained for seconds elapsed between collision and maximum compression is indeed close to one-eighth of a cycle. The assumption that the influence of the spring can be ignored during the collision is justified because the force exerted by the spring is small relative to the force of the impact: At a compression of 15 mm, the magnitude of the force exerted by the spring is $(14\ \text{N/m})(0.015\ \text{m}) = 0.21$ N. The force of impact is given by the time rate of change in the cart's momentum, $\Delta\vec{p}/\Delta t$. The magnitude of the momentum change is $\Delta p = m\Delta v = (0.50\ \text{kg})(0.10\ \text{m/s}) = 0.050\ \text{kg}\cdot\text{m/s}$. If the collision takes place in, say, 20 ms, the magnitude of the force of impact is $(0.050\ \text{kg}\cdot\text{m/s})/(0.020\ \text{s}) = 2.5$ N, which is more than 10 times greater than the magnitude of the force exerted by the spring.

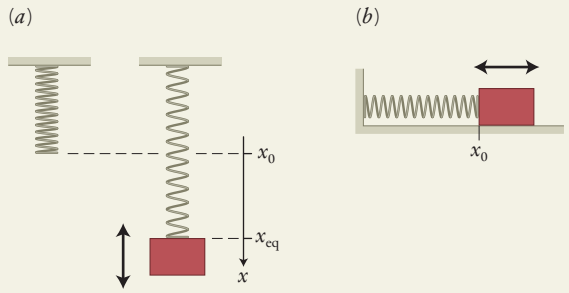
QUANTITATIVE TOOLS

15.13 Suppose cart 2 were not removed from the track of Figure 15.27 immediately after the collision but instead were left stationary at the collision point. (*a*) At what instant do the two carts collide again? (*b*) Describe the motion of the two carts after this second collision.

EXAMPLE 15.5 Vertical oscillations

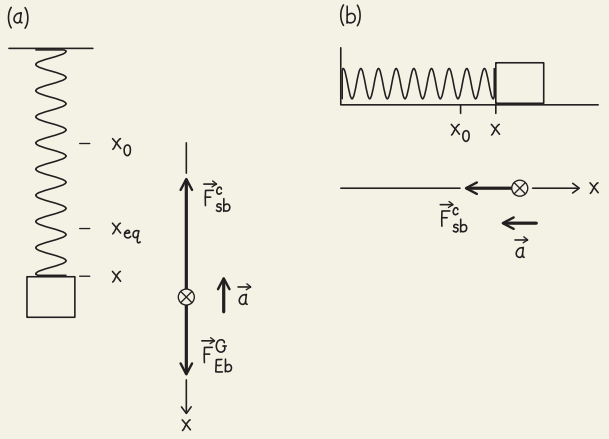
A block of mass $m = 0.50$ kg is suspended from a spring of spring constant $k = 100$ N/m. (*a*) How far below the end of the relaxed spring at x_0 is the equilibrium position x_{eq} of the suspended block (**Figure 15.29a**)? (*b*) Is the frequency f with which the block oscillates about this equilibrium position x_{eq} greater than, smaller than, or equal to that of an identical system that oscillates horizontally about x_0 on a surface for which friction can be ignored (Figure 15.29*b*)?

Figure 15.29 Example 15.5.



❶ **GETTING STARTED** I begin by making a free-body diagram for the suspended block, choosing the positive x axis pointing downward. Two forces are exerted on the block: an upward force $\vec{F}^{c}_{sb}$ exerted by the spring and a downward gravitational force $\vec{F}^{G}_{Eb}$ exerted by Earth (**Figure 15.30a**). When the suspended block is in translational equilibrium at x_{eq} (which lies below x_0), the vector sum of these forces must be zero. With the block at x_{eq}, the spring is stretched such that the end attached to the block is also at x_{eq}. When the block is below x_{eq}, the spring is stretched farther, causing the magnitude of $\vec{F}^{c}_{sb}$ to increase, and so now the vector sum of the forces exerted on the block points upward. When the block is above x_{eq} (but below the position x_0 of the end of the spring when it is relaxed), the spring is stretched less than when the block is at x_{eq}, causing the magnitude of $\vec{F}^{c}_{sb}$ to decrease, and so the vector sum of the forces exerted on the block points downward. The vector sum of $\vec{F}^{c}_{sb}$ and $\vec{F}^{G}_{Eb}$ thus serves as a restoring force.

Figure 15.30



I also make a free-body diagram for the horizontal arrangement (Figure 15.30*b*), showing only the horizontal forces (the force of gravity and the normal force exerted by the surface cancel out). I let the positive x axis point to the right. In this case, the restoring force is $\vec{F}^{c}_{sb}$ only.

❷ **DEVISE PLAN** In translational equilibrium, the vector sum of the forces exerted on the suspended block is zero, and so I can determine the magnitude of the force exerted by the spring. I can then use Hooke's law (Eq. 8.20) to determine the distance between the equilibrium position and x_0. To compare the oscillation frequencies of the two systems, I should write the simple harmonic oscillator equation for each system in the form given by Eq. 15.12.

❸ **EXECUTE PLAN** (*a*) For translational equilibrium, I have

$$\Sigma F_x = F^{c}_{sbx} + F^{G}_{Ebx} = -k(x_{eq} - x_0) + mg = 0,$$

where $x_{eq} - x_0$ is the displacement of the spring's end from its relaxed position. Solving for $x_{eq} - x_0$, I obtain

$$x_{eq} - x_0 = \frac{mg}{k} = \frac{(0.50\text{ kg})(9.8\text{ m/s}^2)}{100\text{ N/m}} = 0.049\text{ m}. \checkmark$$

(*b*) For the horizontal arrangement, I have in the situation depicted in my sketch (Figure 15.30*b*)

$$\Sigma F_x = F^{c}_{sbx} = -k(x - x_0), \qquad (1)$$

so if I let the origin of my x axis be at the position of the end of the relaxed spring, $x_0 = 0$, the rightmost factor in Eq. 1 reduces to $-kx$ and thus $\Sigma F_x = -kx$. Next I turn to the vertical arrangement. In the position illustrated in Figure 15.30*a*, the x component of the upward force exerted by the spring is

$$F^{c}_{sbx} = -k(x - x_0) = -k(x - x_{eq}) - k(x_{eq} - x_0). \quad (2)$$

From part *a* I know that $k(x_{eq} - x_0)$ is equal to mg. The x component of the vector sum of the forces exerted on the block at position x is thus

$$\Sigma F_x = F^{G}_{Ebx} + F^{c}_{sbx} = mg - k(x - x_{eq}) - mg$$

$$= -k(x - x_{eq}).$$

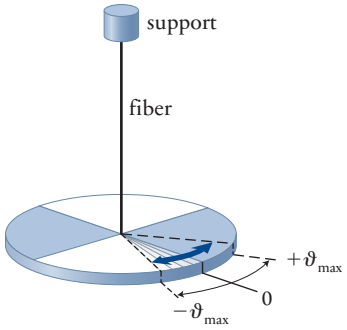
If, as usual, I let the origin be at the equilibrium position, $x_{eq} = 0$, then this result for ΣF_x is identical to Eq. 1. Comparing these results to Eq. 15.12, I see that in both cases $k = m\omega^2$ and so the oscillation frequencies $f = \omega/2\pi$ of the two systems are the same: $f_{vert} = f_{hor}$. ✔

❹ **EVALUATE RESULT** The two oscillations take place about different equilibrium positions (x_0 for the horizontal case, x_{eq} for the vertical case), but the effect of the combined gravitational and elastic forces in the vertical arrangement is the same as that of just the elastic force in the horizontal arrangement because the force exerted by the spring is linear in the displacement.

15.14 Convince yourself that the argument presented in the solution to Example 15.5 is also valid for displacements of the block above x_{eq}.

15.7 Restoring torques

Figure 15.31 Torsional oscillator.



Some simple harmonic oscillators involve rotational rather than translational displacements. As an example, consider the disk suspended from a thin fiber shown in **Figure 15.31**. Such a setup is called a *torsional oscillator*. If the disk is rotated in the horizontal plane over an angle ϑ_{max}, the fiber is twisted, and elastic potential energy is stored in the fiber. When the disk is released, it oscillates as energy is converted back and forth between the elastic potential energy in the fiber and the rotational kinetic energy of the disk. Translational oscillations, as in our simple pendulum and block-spring setup, are due to the interplay between a restoring force and inertia; rotational oscillations are due to an interplay between a restoring torque and rotational inertia.

For a quantitative description of the disk's oscillation, we use the rotational form of the equation of motion for the disk, which relates the sum of the torques caused by the forces acting on the disk to its rotational acceleration α_ϑ (Eq. 12.10):

$$\Sigma\tau_\vartheta = I\alpha_\vartheta, \tag{15.24}$$

where I is the rotational inertia of the disk.

To complete the description, we must relate the torque to the rotational displacement—just as we related force to translational displacement for translational oscillations. In the case of the rotating disk, the torque is caused by the twisted fiber: A clockwise rotation of the fiber causes a counterclockwise torque, and vice versa; greater rotational displacements of the disk cause greater torques. For small rotational displacements, we might expect that the relationship between rotational displacement and torque is linear:

$$\tau_\vartheta = -\kappa(\vartheta - \vartheta_0). \tag{15.25}$$

The minus sign indicates that the direction of the torque is opposite the direction of the rotational displacement: The torque tends to rotate the disk back toward its equilibrium position. This rotational Form of Hooke's law is found to hold for a wide variety of materials. The constant κ (the Greek letter kappa), called the *torsional constant*, depends on the properties of the material being twisted. Just as stiff springs have large spring constants, stiff materials have large torsional constants.

If we let the rotational position at equilibrium be zero, $\vartheta_0 = 0$, we can combine Eqs. 15.24 and 15.25, yielding

$$I\alpha_\vartheta = I\frac{d^2\vartheta}{dt^2} = -\kappa\vartheta \tag{15.26}$$

or

$$\frac{d^2\vartheta}{dt^2} = -\frac{\kappa}{I}\vartheta. \tag{15.27}$$

Notice the similarity between this equation and two earlier ones: Eq. 15.10 for a generic harmonic oscillator and Eq. 15.21 for a mass oscillating on a spring. The meaning of the symbols is different, but the mathematical form is identical. The disk therefore executes simple harmonic motion, and the rotational position $\vartheta(t)$ is a sinusoidal function of time.

To find the position as a function of time, we exploit the correspondence between Eqs. 15.21 and 15.27, substituting rotational displacement ϑ for translational displacement x in Eq. 15.23 and rotational inertia I for mass m and the torsional constant κ for the spring constant k in Eq. 15.22:

$$\vartheta = \vartheta_{\max} \sin(\omega t + \phi_i) \tag{15.28}$$

$$\omega = \sqrt{\frac{\kappa}{I}} \quad \text{(torsional oscillator)}, \tag{15.29}$$

where $\vartheta_{\max}$ in Eq. 15.28 is the maximum rotational displacement (the amplitude of the rotational oscillation). Note that the ω in Eq. 15.29 represents the angular frequency of the torsional oscillator, not its rotational speed.

15.15 For the torsional oscillator shown in Figure 15.31, what effect, if any, does a decrease in the radius of the disk have on the oscillation frequency f? Assume the disk's mass is kept the same.

The pendulum is another example of a rotational oscillator. For a pendulum, however, the restoring torque is not caused by an elastic force, and so we cannot use Eq. 15.25. Instead, we must write an expression for the restoring torque due to gravity and relate the magnitude of this torque to the rotational displacement of the pendulum. Consider an object suspended about a rotation axis located a distance ℓ_{cm} from the object's center of mass, as illustrated in **Figure 15.32a**. The restoring torque is provided by the component of the force of gravity perpendicular to ℓ_{cm}: $F^G_{Eo\perp} = -mg \sin\vartheta$ (Figure 15.32*b*). The lever arm of this force relative to the rotation axis is ℓ_{cm}, and so for a rotational displacement ϑ, the torque caused by the force of gravity about the axis is

$$\tau_\vartheta = -\ell_{cm}(mg \sin\vartheta), \tag{15.30}$$

where m is the object's mass. For small rotational displacements, $\sin\vartheta \approx \vartheta$, and so once again we obtain a restoring torque that is linearly proportional to the rotational displacement:

$$\tau_\vartheta = -(m\ell_{cm}\, g)\vartheta. \tag{15.31}$$

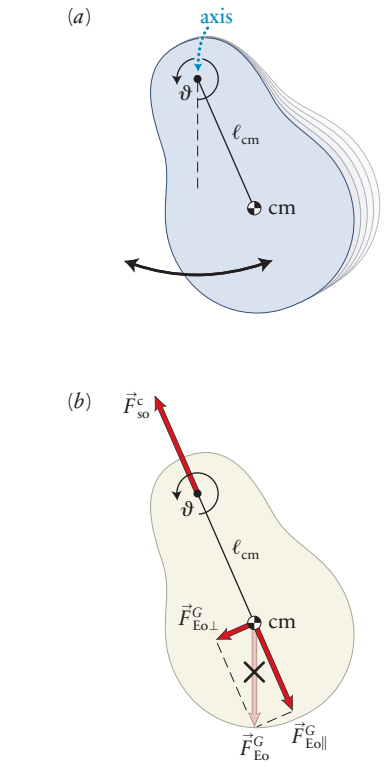
After $\tau_\vartheta = I\alpha_\vartheta = Id^2\vartheta/dt^2$ is substituted for the left side of Eq. 15.31, we get

$$\frac{d^2\vartheta}{dt^2} = -\frac{m\ell_{cm}\, g}{I}\vartheta, \tag{15.32}$$

and so the angular frequency of oscillation is

$$\omega = \sqrt{\frac{m\ell_{cm}\, g}{I}} \quad \text{(pendulum)}. \tag{15.33}$$

Figure 15.32 (*a*) Oscillating pendulum. (*b*) Extended free-body diagram for the pendulum.

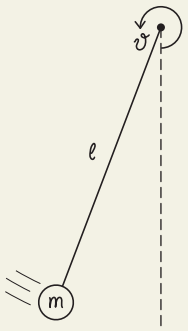


EXAMPLE 15.6 The simple pendulum

Suppose a simple pendulum consisting of a bob of mass m suspended from a string of length ℓ is pulled back and released. What is the period of oscillation of the bob?

1 GETTING STARTED I begin by making a sketch of the simple pendulum (**Figure 15.33**), indicating the equilibrium position by a vertical dashed line.

Figure 15.33



2 DEVISE PLAN The period of the pendulum is related to the angular frequency (Eq. 15.1). To calculate the angular frequency, I can use Eq. 15.33 with $\ell_{cm} = \ell$. If I treat the bob as a particle, I can use Eq. 11.30, $I = mr^2$, with $r = \ell$ to calculate the bob's rotational inertia about the point of suspension.

3 EXECUTE PLAN Substituting the bob's rotational inertia into Eq. 15.33, I get

$$\omega = \sqrt{\frac{m\ell g}{m\ell^2}} = \sqrt{\frac{g}{\ell}},$$

so, from Eq. 15.1, $\omega = 2\pi/T$, I obtain

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{\ell}{g}}. \checkmark$$

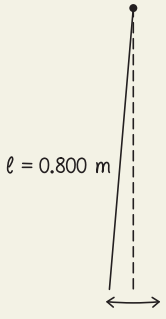
4 EVALUATE RESULT My result says that T is independent of the mass m of the bob, in agreement with what is stated in Section 15.4. Increasing g decreases the period as it should: A greater g means a greater restoring force, and so the bob is pulled back to the equilibrium position faster. It also makes sense that increasing ℓ increases the period: As my sketch shows, for greater ℓ the bob has to move a greater distance to return to the equilibrium position.

EXAMPLE 15.7 Measuring *g*

The oscillations of a thin rod can be used to determine the value of the acceleration due to gravity. A rod that is 0.800 m long and suspended from one end is observed to complete 100 oscillations in 147 s. What is the value of g at the location of this experiment?

1 GETTING STARTED I begin by making a sketch of the situation. To simplify the problem, I assume that the rod pivots about the fixed end (**Figure 15.34**).

Figure 15.34



2 DEVISE PLAN I calculated the rotational inertia of a thin rod about one end in Example 11.11: $I = \frac{1}{3}m\ell^2$, where m is the rod's mass and ℓ its length. I can use this equation and Eq. 15.33 to solve for g, noting that ℓ_{cm} in Eq. 15.33 is half the length of the rod $\left(\ell_{cm} = \frac{1}{2}\ell\right)$.

3 EXECUTE PLAN Substituting the rod's rotational inertia and the center-of-mass distance $\ell_{cm} = \frac{1}{2}\ell$ into Eq. 15.33, I get

$$\omega = \sqrt{\frac{m\frac{1}{2}\ell g}{\frac{1}{3}m\ell^2}} = \sqrt{\frac{3g}{2\ell}}.$$

Solving for g yields $g = \frac{2}{3}\omega^2\ell$. The period of the rod is $(147\text{ s})/(100) = 1.47\text{ s}$, and the angular frequency is $\omega = 2\pi f = 2\pi(1/T) = 2\pi/(1.47\text{ s}) = 4.27\text{ s}^{-1}$. Substituting this information into the expression for g yields $g = \frac{2}{3}(4.27\text{ s}^{-1})^2(0.800\text{ m}) = 9.74\text{ m/s}^2$. ✔

4 EVALUATE RESULT I note that the rod's mass does not appear in my expression for ω. Because ω is proportional to $1/T$, this again confirms what Section 15.4 says about the period of a pendulum being independent of the pendulum's mass because the restoring force is proportional to the pendulum's mass. The value I obtained, 9.74 m/s^2, is reasonably close to the acceleration due to gravity near Earth's surface, giving me confidence in my solution.

15.16 Can you use the experiment described in Example 15.7 to measure the gravitational constant G?

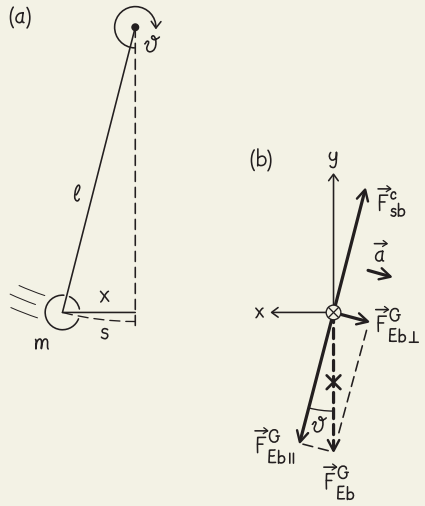
QUANTITATIVE TOOLS

EXAMPLE 15.8 The simple pendulum as a one-dimensional oscillator

The angular frequency ω of a simple pendulum can be calculated by treating the pendulum as a one-dimensional oscillator. In Section 15.4, we used this approach to analyze the restoring force exerted on a pendulum, considering the effect of the force of gravity on the horizontal displacement of the bob and ignoring the slight difference between the horizontal displacement x and the arc length s along which the bob swings. Show that this treatment yields for ω the same expression I obtained in Example 15.6, $\omega = \sqrt{g/\ell}$.

❶ **GETTING STARTED** I make a sketch of the pendulum, showing the bob displaced a horizontal distance x from its equilibrium position, the distance I must concentrate on in this problem (**Figure 15.35*a***). I also draw a free-body diagram for the bob (Figure 15.35*b*), showing $\vec{F}^G_{Eb\parallel}$ and $\vec{F}^G_{Eb\perp}$, with the component vectors of $\vec{F}^G_{Eb}$ parallel to and perpendicular to the string in its displaced orientation.

Figure 15.35



❷ **DEVISE PLAN** The component of the gravitational force perpendicular to the string in its displaced orientation, $\vec{F}^G_{Eb\perp}$, provides the restoring force. For small displacements, this component is nearly parallel to the x axis, and so I can take $F^G_{Eb\perp}$ to be the restoring force in the x direction. This expression will permit me to use the simple harmonic oscillator equation (Eq. 15.10) to obtain an express for ω.

❸ **EXECUTE PLAN** From my free-body diagram, I see that

$$F^G_{Eb\perp} = -mg\sin\theta = -mg\frac{x}{\ell},$$

The x component of the force exerted on the bob is equal to $F^G_{Eb\perp}\cos\theta$. In the small-angle approximation, $F^G_{Eb\perp}\cos\theta \approx F^G_{Eb\perp}$. Thus I can write Eq. 15.11 as

$$ma_x \approx -\frac{mg}{\ell}x,$$

or

$$a_x \equiv \frac{d^2x}{dt^2} \approx -\frac{g}{\ell}x.$$

Comparing this result with Eq. 15.10, $d^2x/dt^2 = -\omega^2 x$, tells me that $-\omega^2 x = -(g/\ell)x$, an equality that yields the same angular frequency of oscillation as in Example 15.6: $\omega = \sqrt{g/\ell}$. ✔

❹ **EVALUATE RESULT** Given that my result is the same as that obtained using a different method, I feel confident that it is correct.

15.17 Imagine placing a pendulum in an elevator. While the elevator accelerates upward, is the frequency f of the pendulum greater than, smaller than, or equal to the frequency when the elevator is at rest?

15.8 Damped oscillations

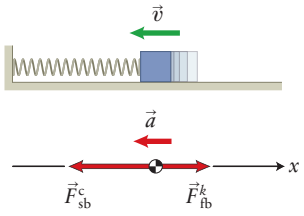
Up to this point, we have ignored any dissipation of energy in simple harmonic motion. Mechanical oscillations always involve some friction, however, and so the energy of the oscillation is slowly converted to thermal energy. As this conversion takes place, the oscillating object slows down. The general term for this slowing down because of energy dissipation is *damping*, and the system is said to execute a **damped oscillation.**

Consider, for example, the two systems in **Figure 15.36**. Figure 15.36*a* shows a block oscillating under the action of a spring. In Figure 15.36*b* the block is replaced by a cart with a vane on top. As the block oscillates, it rubs against the floor, and its motion is slowed down by friction with the floor. The motion of the cart is hindered by air drag caused by the vane.

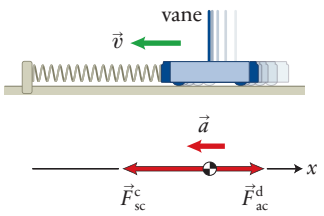
The figure also shows the free-body diagrams for the block and the cart. Each object is subject to two forces in the horizontal direction: a restoring force exerted by the spring and a *damping force*—a force due to friction or drag that dissipates the energy of the oscillation. Damping forces can take

Figure 15.36 Damped oscillating systems.

(*a*) Oscillating block is slowed by friction



(*b*) Oscillating cart is slowed by air drag on vane



QUANTITATIVE TOOLS

many forms. For the oscillating block rubbing on the floor, the magnitude of the force of kinetic friction is $F_{\text{fb}}^{k} = \mu_k(mg)$, with μ_k the coefficient of kinetic friction. The direction of this force is always opposite the direction of motion (that is, in the direction opposite the direction of the velocity vector).

Drag forces exerted by air (or other gases) or liquids at low speeds tend to be proportional to the velocity of the object: The faster you ride your bicycle, for instance, the greater the drag force exerted on you by the air. So

$$\vec{F}_{\text{ac}}^{\text{d}} = -b\vec{v} \quad \text{(low speed)}, \tag{15.34}$$

where the coefficient b, called the *damping coefficient,* depends on the shape of the object on which the drag force acts and on the properties of the gas or liquid exerting the force. The SI units of the damping coefficient are kg/s. The minus sign tells you that the direction of the drag force is opposite the direction of the velocity.

For most oscillators, velocity-dependent drag forces are the main cause of energy dissipation. In the presence of drag, the equation of motion for the cart in Figure 15.36 becomes

$$\Sigma F_x = F_{\text{sc}x}^{\text{c}} + F_{\text{ac}x}^{\text{d}} = -kx - bv_x = ma_x, \tag{15.35}$$

where k is the spring constant and m the mass of the cart. When we write the x components of the velocity and the acceleration as time derivatives of x, Eq. 15.35 can be written in the form

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = 0. \tag{15.36}$$

As long as the damping is not too great, the solution to this differential equation is a sinusoidally varying function whose amplitude decreases with time. The solution, given here without proof, is

$$x(t) = Ae^{-bt/2m}\sin(\omega_{\text{d}}t + \phi_{\text{i}}), \tag{15.37}$$

where the angular frequency ω_{d} of the damped oscillator is given by

$$\omega_{\text{d}} = \sqrt{\frac{k}{m} - \frac{b^2}{4m^2}} = \sqrt{\omega^2 - \left(\frac{b}{2m}\right)^2}. \tag{15.38}$$

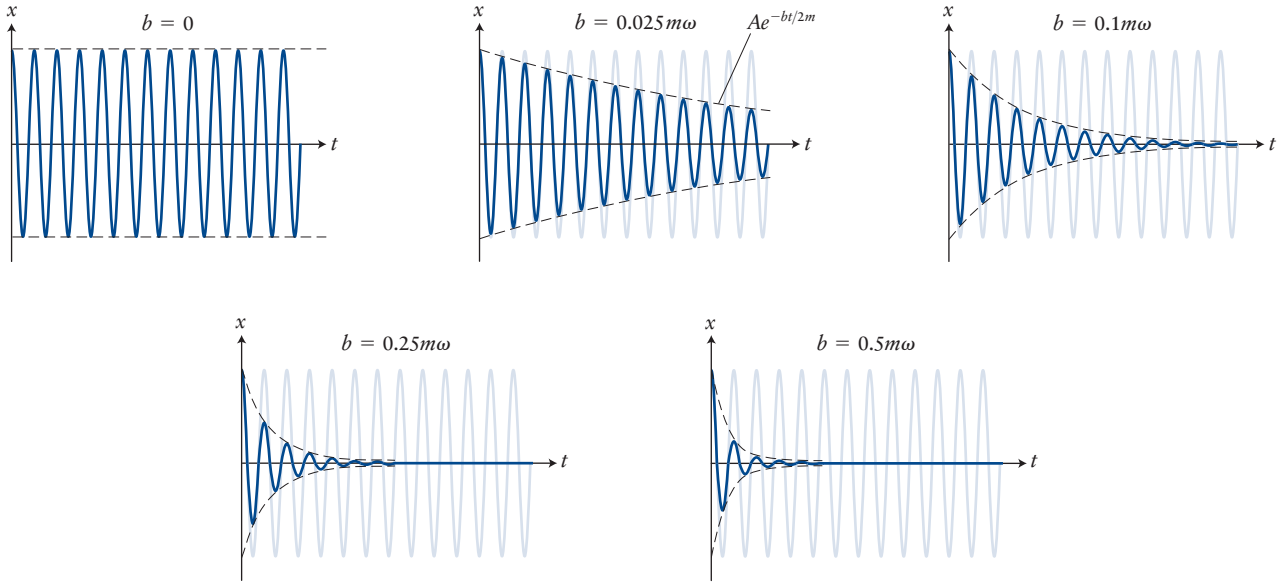
Figure 15.37 shows oscillations for various values of the damping coefficient b. For $b = 0$, there is no damping, which means $\omega_{\text{d}} = \omega$ and $e^{-bt/2m} = 1$. In this case, Eq. 15.37 is identical to Eq. 15.6, the equation for the position of an undamped simple harmonic oscillator. For light damping (small b), the second term in the square root in Eq. 15.38 is negligible, and the damping does not affect the angular frequency of the oscillation: $\omega_{\text{d}} \approx \omega_0$. As Figure 15.37 shows, however, damping decreases the amplitude of the oscillation with each swing. Mathematically, this is expressed by the exponential (see Appendix B) factor $e^{-bt/2m}$ in Eq. 15.37, which makes the amplitude at any given instant equal to $x_{\text{max}} = Ae^{-bt/2m}$, where x_{max} is the time-dependent amplitude of the damped oscillator and A is the (constant) amplitude of the undamped system. In the figure, this decreasing amplitude x_{max} is indicated by the dashed exponential curves, called the *envelope* of the oscillation. The stronger the damping, the faster the amplitude decreases. For very large damping, oscillations cease to occur and the system just returns to the equilibrium position. Such a system is said to be *overdamped*.

The ratio m/b has units of time, and it is customary to call this ratio the **time constant** $\tau \equiv m/b$. (Even though the symbol for the time constant is the same as that for torque, there is no relationship between the two.) With this simplification, the decreasing amplitude can be written as

$$x_{\text{max}}(t) = Ae^{-t/2\tau}. \tag{15.39}$$

QUANTITATIVE TOOLS

Figure 15.37 Damped simple harmonic motion for various values of the damping coefficient b not so great as to prevent oscillations.



The mechanical energy of the oscillator, which is proportional to the square of the amplitude (Eq. 15.17), thus decreases as

$$E(t) = \tfrac{1}{2}m\omega^2 x_{\max}^2 = (\tfrac{1}{2}m\omega^2 A^2)e^{-t/\tau} = E_0 e^{-t/\tau}, \qquad (15.40)$$

where $E_0 = \frac{1}{2}m\omega^2 A_0^2$ is the initial mechanical energy. Equation 15.40 shows that after a time interval equal to the time constant, so that $t = \tau$, the energy has decreased to $e^{-\tau/\tau} = e^{-1} \approx 0.37$ of its initial value; only about one-third of the initial energy remains. During each time interval of duration τ, the energy again decreases by the factor e^{-1}. After two time constants, for instance, only about 10% of the initial energy remains.

Because oscillation periods vary widely, it is often convenient to express the time constant τ in terms of the period T so that we know after how many *cycles* (rather than seconds) the energy has decreased by a certain factor. In terms of cycles, the decay of the oscillation can be characterized by the *quality factor,* defined as

$$Q \equiv \omega\tau = 2\pi\frac{\tau}{T}. \qquad (15.41)$$

The quality factor is a unitless ratio similar to the rotational coordinate ϑ—if $Q = 2\pi$, then $\tau = T$, and so the energy decays by a factor of e^{-1} in one cycle. When $Q = 1000(2\pi)$, it takes 1000 cycles for the energy to decay by a factor of e^{-1}. A large Q implies that $\omega_d \approx \omega_0$. For mechanical oscillators with low friction (piano strings, tuning forks), Q can be as large as a few thousand, meaning the oscillator executes thousands of oscillations before its energy decreases by a factor of e^{-1}. A bell with a large Q rings for a longer time interval than one with a small Q.

15.18 A tuning fork that sounds the tone musicians call middle C oscillates at frequency $f = 262$ Hz. If the amplitude of the fork's oscillation decreases by a factor of 3 in 4.0 s, what are (*a*) the time constant of the oscillation and (*b*) the quality factor?

Chapter Summary

Fundamental characteristics of periodic motion (Sections 15.1, 15.5)

Concepts **Periodic motion** is any motion that repeats itself at regular time intervals. **Oscillation** (or **vibration**) is back-and-forth periodic motion.

The *period* T is the minimum time interval in which periodic motion repeats, and the **amplitude** A of the motion is the magnitude of the maximum displacement of the moving object from its equilibrium position.

Quantitative Tools The **frequency** f of periodic motion is the number of cycles per second and is defined as

$$f \equiv \frac{1}{T}. \tag{15.2}$$

The SI unit of frequency is the **hertz** (Hz):

$$1 \text{ Hz} \equiv 1 \text{ s}^{-1}. \tag{15.3}$$

Simple harmonic motion (Sections 15.2, 15.4, 15.5)

Concepts **Simple harmonic motion** is periodic motion in which the displacement of a system from its equilibrium position varies sinusoidally with time. A system moving in this way is called a *simple harmonic oscillator*.

A restoring force that is linearly proportional to displacement tends to return a simple harmonic oscillator to its equilibrium position. For small displacements, restoring forces are generally proportional to the displacement and therefore cause objects to execute simple harmonic motion about any stable equilibrium position.

A **phasor** is a rotating arrow whose component on a vertical axis traces out simple harmonic motion. The **reference circle** is the circle traced out by the tip of the phasor, and the length of the phasor is equal to the amplitude A of the simple harmonic motion.

Quantitative Tools The **angular frequency** ω of oscillation is equal to the rotational speed of a rotating phasor whose vertical component oscillates with a frequency f. Angular frequency is measured in s^{-1} and is related to the frequency (measured in Hz) by

$$\omega = 2\pi f. \tag{15.4}$$

For a simple harmonic oscillator of amplitude A, the displacement x as a function of time is

$$x(t) = A \sin(\omega t + \phi_i), \tag{15.6}$$

where the sine argument is the **phase** $\phi(t)$ of the periodic motion,

$$\phi(t) = \omega t + \phi_i, \tag{15.5}$$

and ϕ_i is the *initial phase* at $t = 0$.

The x components of the velocity and acceleration of a simple harmonic oscillator are

$$v_x \equiv \frac{dx}{dt} = \omega A \cos(\omega t + \phi_i) \tag{15.7}$$

$$a_x \equiv \frac{d^2x}{dt^2} = -\omega^2 A \sin(\omega t + \phi_i). \tag{15.8}$$

Any object that undergoes simple harmonic motion obeys the **simple harmonic oscillator equation:**

$$\frac{d^2x}{dt^2} = -\omega^2 x, \tag{15.10}$$

where x is the object's displacement from its equilibrium position.

The mechanical energy E of an object of mass m undergoing simple harmonic motion is

$$E = \tfrac{1}{2} m\omega^2 A^2. \tag{15.17}$$

Fourier series (Section 15.3)

Concepts **Fourier's theorem** says that any periodic function with period T can be written as a sum of sinusoidal simple harmonic functions of frequency $f_n = n/T$, where n is an integer. The $n = 1$ term is the **fundamental frequency** or **first harmonic,** and the other components are **higher harmonics.**

Oscillating springs (Section 15.6)

Quantitative Tools For an object attached to a light spring of spring constant k, the simple harmonic oscillator equation takes the form

$$\frac{d^2x}{dt^2} = -\frac{k}{m}x, \tag{15.21}$$

and the angular frequency of the oscillation is

$$\omega = +\sqrt{\frac{k}{m}}. \tag{15.22}$$

The motion of the object is described by

$$x(t) = A\sin\left(\sqrt{\frac{k}{m}}t + \phi_i\right). \tag{15.23}$$

Rotational oscillations (Section 15.7)

Concepts A horizontal disk suspended at its center by a thin fiber forms a type of *torsional oscillator*.

A pendulum is any object that swings about a pivot. A *simple pendulum* consists of a small object (the bob) attached to a very light wire or rod.

Quantitative Tools If a torsional oscillator of rotational inertia I is twisted through a small angle from its equilibrium position ϑ_0 to position ϑ, the restoring torque τ_ϑ is

$$\tau_\vartheta = -\kappa(\vartheta - \vartheta_0), \tag{15.25}$$

where κ is the *torsional constant*. When $\vartheta_0 = 0$, the simple harmonic oscillator equation for the torsional oscillator is

$$\frac{d^2\vartheta}{dt^2} = -\frac{\kappa}{I}\vartheta. \tag{15.27}$$

The rotational position ϑ of the torsional oscillator at instant t is given by

$$\vartheta = \vartheta_{max}\sin(\omega t + \phi_i), \tag{15.28}$$

where ϑ_{max} is the maximum rotational displacement and

$$\omega = \sqrt{\frac{\kappa}{I}}. \tag{15.29}$$

For small rotational displacements, the simple harmonic oscillator equation of a pendulum is

$$\frac{d^2\vartheta}{dt^2} = -\frac{m\ell_{cm}g}{I}\vartheta \tag{15.32}$$

and its angular frequency is

$$\omega = \sqrt{\frac{m\ell_{cm}g}{I}}, \tag{15.33}$$

where ℓ_{cm} is the distance from the center of mass of the pendulum to the pivot.

The period of a simple pendulum is $T = 2\pi\sqrt{\frac{\ell}{g}}$.

Damped oscillations (Section 15.8)

Concepts In **damped oscillation,** the amplitude decreases over time due to energy dissipation. The cause of the dissipation is a *damping force* due to friction, air drag, or water drag.

A damped oscillator that has a large *quality factor* Q keeps oscillating for many periods.

Quantitative Tools At low speeds, the damping force $\vec{F}^{\,\mathrm{d}}_{\mathrm{ao}}$ tends to be proportional to the velocity of the object:

$$\vec{F}^{\,\mathrm{d}}_{\mathrm{ao}} = -b\vec{v}, \tag{15.34}$$

where b, the *damping coefficient,* has units of kilograms per second.

For small damping, the position $x(t)$ of a damped spring is

$$x(t) = Ae^{-bt/2m}\sin(\omega_{\mathrm{d}}t + \phi_{\mathrm{i}}), \tag{15.37}$$

and its angular frequency ω_{d} is

$$\omega_{\mathrm{d}} = \sqrt{\frac{k}{m} - \frac{b^2}{4m^2}} = \sqrt{\omega^2 - \left(\frac{b}{2m}\right)^2}. \tag{15.38}$$

The **time constant** τ for a damped system is $\tau = m/b$. In one time constant, the energy of a damped simple harmonic oscillator is reduced by a factor of $1/e$.

The amplitude $x_{\max}(t)$ and energy $E(t)$ of a damped oscillation of initial amplitude A and initial energy E_0 decrease exponentially with time:

$$x_{\max}(t) = Ae^{-t/2\tau} \tag{15.39}$$

$$E(t) = E_0e^{-t/\tau}. \tag{15.40}$$

The *quality factor* Q of a damped system is

$$Q = 2\pi\frac{\tau}{T}. \tag{15.41}$$

Questions and Problems

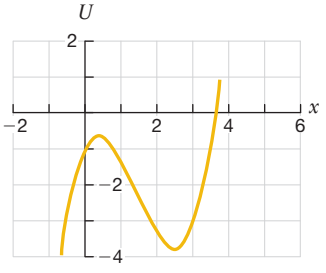
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

15.1 Periodic motion and energy

1. If you photograph a flock of birds taking off from ground level, some of the wings will be blurred in the photograph even though the bodies are in focus. In which wing positions are the wings least blurry? •
2. Plot kinetic energy and potential energy as a function of time for the cart in Figure 15.2. •
3. Based on the energy diagrams in Figure 15.2, sketch a graph showing velocity as a function of position for the cart represented in the figure. ••
4. (*a*) At what displacement (expressed as a fraction of the amplitude of the motion) is the kinetic energy of the cart in Figure 15.2 half its maximum value? (*b*) What is the velocity (expressed as a fraction of the maximum velocity) at this displacement? ••
5. Figure P15.5 shows a graph of the potential energy U of a moving object as a function of its position x. What is the maximum range of x values possible for periodic motion in this system? ••

Figure P15.5

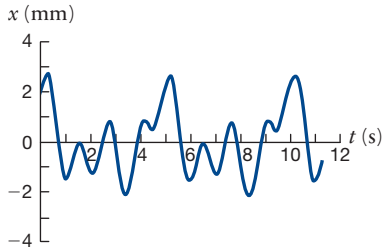


15.2 Simple harmonic motion

6. An object attached to one end of a spring makes 25 complete oscillations in 5.0 seconds. What are its period and frequency? •

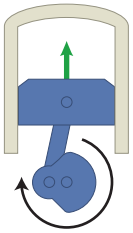
7. What is the period of the motion represented by the curve in Figure P15.7? •

Figure P15.7



8. (*a*) Explain why the motion of the engine piston in Figure P15.8 is only approximately simple harmonic motion. (*b*) What must you do to the length of the connecting rod if you want to make the motion more nearly simple harmonic motion? ••

Figure P15.8

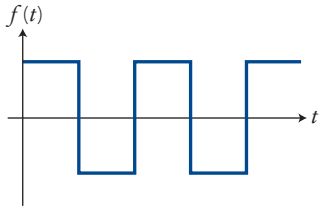


9. Is the typical up-and-down motion of a yo-yo periodic motion, assuming no energy is dissipated? Is it simple harmonic motion? ••
10. A highly elastic ball is dropped from a height of 2.0 m onto a hard surface. Assume the collision is elastic and no energy is lost to air friction. (*a*) Show that the ball's motion after it hits the surface is periodic. (*b*) Determine the period of the motion. (*c*) Is it simple harmonic motion? Why or why not? ••

15.3 Fourier's theorem

11. Fourier analysis of a particular periodic function includes the harmonic frequencies 889 Hz, 1143 Hz, and 1270 Hz. What is the fundamental frequency f_1, assuming that $f_1 > 100$ Hz? •
12. Even when a piano and a trumpet play the same note (which means the sounds have the same frequency f), the two instruments sound completely different. What is it about the sound from the instruments that accounts for this difference? ••
13. Estimate the number of harmonics (counting the fundamental frequency) that contribute significantly in the Fourier analysis of the spectrum in Figure P15.7. ••
14. Suppose you create a periodic function by adding together only sine functions whose amplitudes decrease like $1/n^2$, where n is any odd integer multiple of the fundamental frequency, and where every other term is subtracted rather than added. (*a*) What general properties would this periodic function exhibit? (*b*) Describe the shape of the function obtained by adding only the four terms of frequencies f, $3f$, $5f$, and $7f$. (Hint: Construct the graph using superposition.) •••
15. What harmonic series is required to produce a "square wave" of amplitude 1, as illustrated in Figure P15.15? •••

Figure P15.15



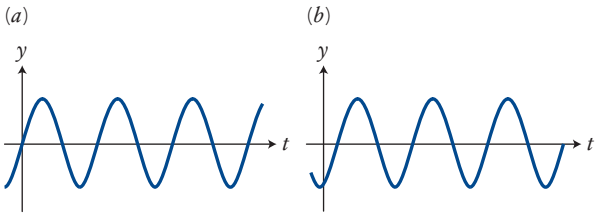
15.4 Restoring forces in simple harmonic motion

16. Which of these forces could result in simple harmonic motion: (*a*) $F(x) = 2x$, (*b*) $F(x) = -2x$, (*c*) $F(x) = -2x^2$, (*d*) $F(x) = 2x^2$, (*e*) $F(x) = -2x + 2$, (*f*) $F(x) = -2(x-2)^2$? •
17. When you dive off a diving board, the board oscillates while you are flying toward the swimming pool. (These oscillations are significantly damped for a good diving board.) Identify the restoring force (or torque) responsible for these oscillations as well as the relevant mass. ••
18. A spring-cart system and a simple pendulum both have the same period near Earth's surface. What happens to the period of each motion when both systems are placed on the surface of the Moon, which has an acceleration due to gravity of 1/6 of that of Earth's? ••
19. Figure 15.12 illustrates the x component of the vector sum of forces on an object for conditions of stable, unstable, and neutral equilibrium. (*a*) Examine this graph carefully. What visible feature differs in the three types of equilibrium? (*b*) Express your result as a set of mathematical relationships that can be used to discriminate stable from unstable or neutral equilibrium. (*c*) Are these relationships general? That is, can you construct a graph showing the x component of the vector sum of forces versus position that meets the requirements for stable equilibrium but violates your mathematical expression? ••
20. The period of a pendulum on the surface of the Earth is 1. It is then brought at a higher altitude. For it to maintain its period, should the length of the pendulum string be made longer or shorter? ••

15.5 Energy of a simple harmonic oscillator

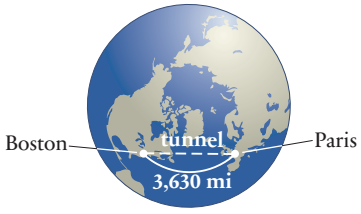
21. Write Eqs. 15.6, 15.7, and 15.8 in terms of the period T of the motion. •
22. Is the initial phase ϕ_i positive, zero, or negative for each oscillation graphed in Figure P15.22? •

Figure P15.22



23. In Figure 15.22, the vertical component of the phasor moves in simple harmonic motion. (*a*) Does the horizontal component also move in simple harmonic motion? (*b*) What is the phase difference $\Delta\phi = \phi_v - \phi_h$ between the vertical and horizontal components? ••
24. On average, Mars takes 687 days to make a complete revolution around the sun. Considering its orbit as nearly circular as seen by a distant observer standing in the plane of the orbit, what is the effective "spring constant" of this simple harmonic motion? ••
25. Consider the reference circle shown in Figure 15.22. (*a*) In terms of the rotational speed ω of the phasor, what is the magnitude of the tangential velocity $\vec{v}$ of the phasor tip? What is the vertical component of this velocity? How does this phasor velocity compare with the velocity of an object moving in simple harmonic motion? (*b*) In terms of the rotational speed ω of the phasor, what is the magnitude of the centripetal acceleration of the phasor tip? What is the vertical component of this acceleration? How does this phasor acceleration compare with the acceleration of an object moving in simple harmonic motion? ••
26. A 0.250-kg mass on a spring has velocity as a function of time given by $v(x, t) = (4.70\ \text{cm/s})\cos[(4.16\ \text{rad/s})t]$. Find the (*a*) the period; (*b*) the amplitude; (*c*) the frequency; and (*d*) the force constant of the spring? ••
27. A toy, initially equilibrium position, attached to a spring undergoes simple harmonic motion horizontally with a period of 0.75 s and amplitude of 4.0 cm. Write an equation describing the toy's motion with all variables identified. •
28. You are standing next to a table and looking down on a record player sitting on the table. Take the spindle (axis of rotation) to be the center of your coordinate system and the y axis to be perpendicular to the side of the player you are standing next to. Long-playing records revolve $33\frac{1}{3}$ times per minute. You put a small blob of clay at the edge of a record that has a radius of 0.15 m, positioning the clay such that it is at its greatest value of y at $t = 0$. (*a*) What is the rotational speed of the clay? (*b*) Write the equation of motion for the y component of the clay's position. ••
29. A 0.200-kg small block is attached to an ideal spring with a spring constant of 316 N/m and is moving in simple harmonic motion on a horizontal frictionless surface. Find the energy of the block if its maximum displacement from the equilibrium position is 16.5 cm. ••
30. The position of a particle undergoing simple harmonic motion is given by $x(t) = a\cos(bt + \pi/3)$, where $a = 8.00$ m and $b = 2.00\ \text{s}^{-1}$. What are the (*a*) amplitude, (*b*) frequency f, and (*c*) period? (*d*) What are the speed and acceleration magnitude of the particle at $t = \pi/2$ s? (*e*) What is its maximum acceleration magnitude, and at what instant after $t = 0$ does the particle first have this acceleration? (*f*) What is its maximum speed, and at what instant after $t = 0$ does it first have this speed? ••
31. For any simple harmonic motion, the position, velocity, and acceleration can all be written as the same type of trigonometric function (sine or cosine) by correctly adding a *relative phase factor*. (*a*) Show that an x component of the position function given by $x(t) = \beta\cos(\omega t + \delta)$ is also a valid general solution for simple harmonic motion. (*b*) Write the x components of the velocity and acceleration associated with this position function. (*c*) Draw the reference circles for the x components of position, velocity, and acceleration, with the phasors correctly drawn for $t = 0$ and the amplitude clearly indicated. •••
32. You get a crazy idea to dig a tunnel through Earth from Boston to Paris, a surface distance of 5850 km, and run a passenger train between the two cities (Figure P15.32). The train would move down the first part of the tunnel under the pull of gravity and then coast upward against gravity. If you assume that the motion will be simple harmonic motion, how long will a one-way trip take? •••

Figure P15.32

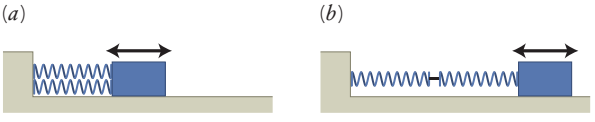


33. You have developed a method in which a paint shaker is used to measure the coefficient of static friction between various objects and a known surface. The shaker oscillates with a fixed amplitude of 50 mm, but you can adjust the frequency of the motion. You have affixed a horizontal tabletop (the known surface) to the shaker so that the tabletop oscillates with it. Then you put an object on the tabletop and increase the frequency until the object begins to slip on the surface. If a frequency $f = 1.85$ Hz is required before a penny positioned on the tabletop starts to slide, what is the coefficient of static friction between penny and tabletop? •••

15.6 Simple harmonic motion and springs

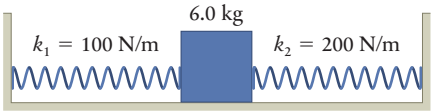
34. Even a spring without a block hanging on the end of it has an oscillation period. Why isn't the period zero? •
35. A vertical spring, on which a block of mass m_1 is hung, oscillates with period T. When replaced with mass m_2, the period is $T/3$. What is the ratio m_1/m_2? ••
36. Two vertical springs, one with spring constant k and the other with spring constant $2k$, each have a ball of mass m hanging from them. Compare the period of oscillation of the two springs undergoing simple harmonic motion. ••
37. A horizontal spring-block system made up of one block and one spring has oscillation frequency f. A second spring, identical to the first, is to be added to the system. (*a*) Does f increase, decrease, or stay the same when the two springs are connected as shown in Figure P15.37*a*? (*b*) What happens to f when the springs are arranged as in Figure P15.37*b*? ••

Figure P15.37



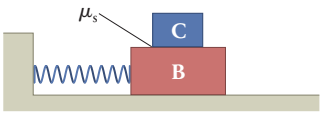
38. An astronaut who recently landed on an unfamiliar planet wants to measure the acceleration due to gravity. He constructs a simple pendulum of length 40.0 cm and then finds that the pendulum makes 35.0 complete swings per minute. Find the value of g on this planet. ••
39. A table outfitted with springs on its feet bounces vertically in simple harmonic motion. A cup of coffee is sitting on the table. Discuss whether any of the following changes to this system could cause the cup to lose contact with the table: (*a*) increase the amplitude, (*b*) increase the period, (*c*) decrease the period, (*d*) decrease the mass of the table, (*e*) increase the mass of the table, (*f*) change the phase of the motion. (*g*) Would your answers be different if you considered not the cup bottom losing contact with the table but rather the coffee flying out of the cup? ••
40. A 5.0-kg cart is attached to a horizontal spring for which the spring constant is 75 N/m. The system is set in motion when the cart is 0.32 m from its equilibrium position, and the initial velocity is 1.5 m/s directed towards the equilibrium position. (*a*) What is the amplitude of the oscillation? (*b*) What is the speed of the cart at the equilibrium position? (*c*) What is the speed of the cart at maximum displacement? (*d*) How would the answers to parts *a*, *b*, and *c* change if the initial direction of the motions is directed away from the cart's equilibrium position? •••
41. A vertical spring-block system with a period of 2.3 s and a mass of 0.35 kg is released 40 mm below its equilibrium position with an initial upward velocity of 0.12 m/s. For this system, determine the (*a*) amplitude, (*b*) angular frequency ω, (*c*) energy, (*d*) spring constant, (*e*) initial phase angle, and (*f*) equation of motion. ••
42. A 5.0-kg object is suspended from the ceiling by a strong spring, which stretches 0.10 m when the object is attached. The object is lifted 0.050 m from this equilibrium position and released. Determine the amplitude and period of the resulting simple harmonic motion. ••
43. A 5.0-kg block is suspended from the ceiling by a strong spring and released to perform simple harmonic motion with a period of 0.50 s. The block is brought to rest, and the length of the spring with the block attached is measured. By how much is this length reduced when the block is removed? ••
44. A 4.0-kg object is suspended from the ceiling by a spring and undergoes simple harmonic motion with an amplitude of 0.50 m. At the highest position in the motion, the spring is at the length that would be its relaxed length if no object were attached. (*a*) Calculate the energy of the system. For the object at its lowest position, its equilibrium position, and its highest position, calculate (*b*) the elastic potential energy of the spring, (*c*) the kinetic energy of the object, and (*d*) the gravitational potential energy of the Earth-object system. ••
45. A 6.0-kg block free to slide on a horizontal surface is anchored to two facing walls by springs (Figure P15.45). Both springs are initially at their relaxed length. The block is then displaced 20 mm to the right and released. (*a*) What is the effective spring constant of the system? (*b*) What is the equation for the position of the block as a function of time, assuming $+x$ to be to the right? Ignore friction. ••

Figure P15.45



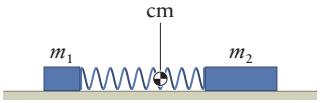
46. Block B in Figure P15.46 is free to slide on the horizontal surface. With block C placed on top of B, the system undergoes simple harmonic motion with an amplitude of 0.10 m. Block B has a speed of 0.24 m/s at a displacement of 0.060 m from its equilibrium position. (*a*) Determine the period of the motion. (*b*) What minimum value for the coefficient of static friction μ_s between B and C is needed if C is never to slip? Ignore any friction between B and the horizontal surface. •••

Figure P15.46



47. After gluing two 0.50-kg blocks together, you determine that you can pull the blocks apart with a force of 20 N applied to one block. You attach the top block to a vertical spring for which the spring constant is 500 N/m. What is the maximum possible amplitude of oscillation for which the glue can keep the blocks together? •••
48. Two blocks, of masses m_1 and m_2, are placed on a horizontal surface and attached to opposite ends of a spring as in Figure P15.48. The blocks are then pushed toward each other to compress the spring. When the blocks are released, (*a*) describe the motion of the center of mass of the system and (*b*) derive an expression for the angular frequency ω of the motion. (*c*) What is ω in the limit $m_2 \gg m_1$? (*d*) What is ω when $m_1 = m_2$? Ignore friction and the mass of the spring. •••

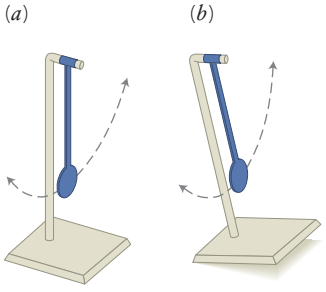
Figure P15.48



15.7 Restoring torques

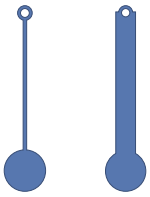
49. You want to build a pendulum clock in which the time interval between the "tick" sound (when the pendulum is swinging one way) and the "tock" sound (when the pendulum is swinging the other way) is 0.50 s. If we assume the pendulum is a simple one, what should its length be? •
50. What happens to the period of the pendulum shown in Figure P15.50*a* when the stand supporting the pendulum is tipped backward as shown in Figure P15.50*b*? •

Figure P15.50



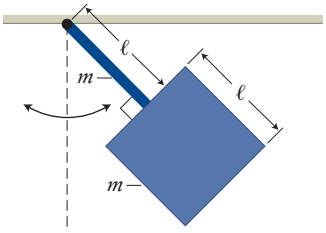
51. Which pendulum in Figure P15.51 has the greater period? ••

Figure P15.51



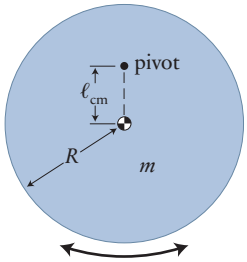
52. A small-angle approximation was used to derive Eq. 15.31, $\tau = -(m\ell g)\vartheta$. (*a*) What constitutes *small* in this context? In other words, how large can ϑ be before it can no longer be called small? (*b*) As a quantitative benchmark, how large does ϑ have to be before $(m\ell g)\vartheta$ deviates by more than 1% from the true value of τ? ••
53. The maximum angle from the vertical reached by a simple pendulum of length ℓ is ϑ_{max}. What is the linear speed v of the bob for any angle ϑ from the vertical? ••
54. The 800-mg balance wheel of a certain clock is made up of a thin metal ring of radius 15 mm connected by spokes of negligible mass to a fine suspension fiber as in Figure 15.31. The back-and-forth twisting of the fiber causes the wheel to move in simple harmonic motion with period T. The clock ticks four times each second, with the interval between ticks equal to $T/2$. (*a*) Determine the rotational inertia I of the balance wheel. (*b*) What is the torsional constant of the fiber that drives the balance wheel? ••
55. A thin 0.100-kg rod that is 250 mm long has a small hole drilled through it 62.5 mm from one end. A metal wire is strung through the hole and fixed horizontally, and the rod is free to rotate about the wire and to oscillate. (*a*) Determine the rod's rotational inertia I about this axis. (*b*) Determine the period of the rod's oscillation. ••
56. A rod of mass m and length ℓ is rigidly connected at a right angle to the midpoint of one face of a uniform cube of mass m and side length ℓ (Figure P15.56). The other end of the rod is connected to a pivot so that the system can oscillate freely. What is the period of small oscillations about the pivot? ••

Figure P15.56



57. A simple pendulum of length 0.30 m has a 0.30-kg bob. At $t = 0$, the bob passes through the lowest position in its motion, and at this instant it has a horizontal speed of 0.25 m/s. (*a*) What is the maximum angular displacement ϑ_{max} away from vertical that the pendulum reaches? (*b*) What is the bob's linear speed v when it has angular displacement $\vartheta_{max}/2$? ••
58. A uniform disk of mass m and radius R lies in a vertical plane and is pivoted about a point a distance ℓ_{cm} from its center of mass (Figure P15.58). When given a small rotational displacement about the pivot, the disk undergoes simple harmonic motion. Determine the period of this motion. ••

Figure P15.58



59. A thin 0.50-kg ring of radius $R = 0.10$ m hangs vertically from a horizontal knife-edge pivot about which the ring can oscillate freely. If the amplitude of the motion is kept small, what is the period? ••
60. If a person is 20% shorter than average, what is the ratio of his walking pace (that is, the frequency f of his motion) to the walking pace of a person of average height? Assume that a person's leg swings like a pendulum and that the angular amplitude of everybody's stride is about the same. •••
61. A pendulum consists of a bob of mass m_{bob} hanging at the end of a light rod of mass m_{rod} and length ℓ. If you ignore the mass of the rod, the period is T_0, and if you use m_{rod} in your calculation, the period is T. How far from the true value T is T_0? Express the error $\Delta T = T - T_0$ in terms of some power of the ratio m_{rod}/m_{bob}. (Hint: Some sort of series expansion is useful here.) •••

15.8 Damped oscillations

62. Verify that the time constant $\tau = m/b$ has units of time for damped oscillations. •
63. As the quality factor Q for a damped oscillation increases but the period T of undamped motion is held constant, what happens to the angular frequency ω_d of the damped motion? •
64. Express, as a function of τ, the time interval needed for the amplitude of an oscillation to reduce to half of its initial value. •
65. A 0.400-kg object is oscillating on a spring for which the force constant is 300 N/m. A damping force that is linearly proportional to the object's velocity is exerted on the system, and the damping coefficient is $b = 5.00$ kg/s. (*a*) Verify that the units given for b are correct. (*b*) What is the frequency f of the oscillation? (*c*) What is the time constant τ? (*d*) What is the quality factor Q? ••
66. A simple pendulum consists of a 1.00-kg bob on a string 1.00 m long. During a time interval of 27.0 s, the maximum angle this pendulum makes with the vertical is found to decrease from 6.00° to 5.40°. Determine the numerical values of the damping coefficient b and the time constant τ. ••
67. A 0.25-kg bob is suspended from a string that is 0.60 m long. When the pendulum is set into small-amplitude oscillation, the amplitude decays to half of its initial value in 35 s. (*a*) What is the time constant τ for the pendulum? (*b*) At what instant is the energy half of its initial value? ••
68. A restoring force of unknown magnitude is exerted on a object that oscillates with a period of 0.50 s. When the object is in an evacuated container, the motion is simple harmonic motion, with an amplitude of 0.10 m. When

air is allowed into the container, the amplitude decreases by 2.0% with each cycle of the oscillation. (*a*) What is the amplitude after 25 cycles? (*b*) What fraction of the initial energy is left 6.3 s after air is admitted? ••

69. A 0.500-kg block oscillates up and down on a vertical spring for which $k = 12.5$ N/m. The initial amplitude of the motion is 0.100 m. (*a*) What is the natural angular frequency ω? (*b*) If the observed angular frequency is $\omega_d = 4.58\ s^{-1}$, what is b? (*c*) Write an equation for the position as a function of time. ••

70. A damped oscillator has a quality factor of 20. (*a*) By what fraction does the energy decrease during each cycle? (*b*) By what percentage does the damped angular frequency ω_d differ from the undamped angular frequency? ••

71. One reference book states that, after a large earthquake, Earth can be treated as an oscillator that has a period of 54 min and a quality factor of 400. (*a*) What percentage of the energy of oscillation is lost to damping forces during each cycle? (*b*) What fraction of the original energy is left after 2.0 days? ••

72. A gong makes a loud noise when struck. The noise gradually gets less and less loud until it fades below the sensitivity of the human ear. The simplest model of how the gong produces the sound we hear treats the gong as a damped harmonic oscillator. The tone we hear is related to the frequency f of the oscillation, and its loudness is proportional to the energy of the oscillation. (*a*) If the loudness drops to 85% of its original value in 4.0 s, what is the time constant of the damped oscillation? (*b*) How long does it take for the sound to be 25% as loud as it was at the start? (*c*) What fraction of the original loudness remains after 1.0 min? ••

73. *Critical damping* is a term used to describe a situation where the damping coefficient b is just great enough that no oscillation occurs. (In critical damping, a moving part is displaced but returns to its equilibrium position quickly without overshooting it.) This is how you want the springs in your car to behave, so that when they flex to take you smoothly over a bump, they don't cause the car to jump up and down, reducing contact with the road.) (*a*) For an object of mass m attached to a spring of spring constant k, what is the value b_{crit} needed for critical damping? (*b*) Describe the motion if $b > b_{crit}$. ••

74. A small 3.0-kg object dropped from the roof of a tall building acquires a terminal speed of 25 m/s. Assume the drag force exerted on the object has the same form as the damping force exerted on a damped oscillator; that is, the force is opposed to the motion, and its magnitude is linearly proportional to the object's speed. An object identical to the dropped one is attached to a vertical spring ($k = 230$ N/m) and set into oscillation in the air with an initial amplitude of 0.20 m. (*a*) What is the quality factor for this oscillation? (*b*) How long does it take for the amplitude to drop to half its initial value? (*c*) How much energy is lost in this time interval? •••

75. An object resting on a table oscillates on the end of a horizontal spring. Because the table is covered with a viscous substance, the motion is damped and the amplitude gets smaller with each cycle. At $t = 1.5$ s, the object is 60 mm from its equilibrium position, and this is the farthest from equilibrium it reaches. The amplitude of the next cycle of the motion is 56 mm, and the object reaches this maximum x value at $t = 2.5$ s. Write the equation for the object's x component of position as a function of time, using numerical values for all constants. •••

Additional Problems

76. Do the turning points in simple harmonic motion have to be equidistant from the equilibrium position? •

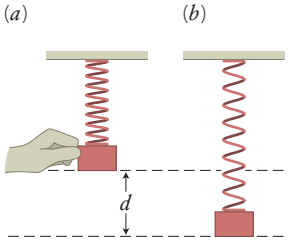
77. A mass attached to a spring oscillates in simple harmonic motion along the x-axis. The limits of its motion are $x = -5$ cm and $x = 45$ cm and it goes from one of these extremes to the other in 0.20 s. Find its (*a*) amplitude; (*b*) frequency; and (*c*) period? ••

78. A pendulum is swinging in a stationary elevator. What happens to the period when the elevator (*a*) accelerates upward, (*b*) travels downward at constant speed, and (*c*) travels downward and gradually slows to a stop? •

79. You measure the oscillation frequency f_{whole} of a vertical block-spring system. You then cut the spring in half, hang the same block from one of the halves, and measure the frequency f_{half}. What is the ratio f_{half}/f_{whole}? ••

80. The hand in Figure P15.80*a* holds a vertical block-spring system so that the spring is compressed. When the hand lets go, the block, of mass m, descends a distance d in a time interval Δt before reversing direction (Figure P15.80*b*). Once the system comes to rest, the block is removed, a block of mass $2m$ is hung in its place, and the experiment is repeated. To what maximum distance does the block descend? ••

Figure P15.80

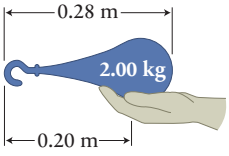


81. Two oscillatory motions are given by $x(t) = A\sin(m\omega t)$ and $y(t) = A\sin(n\omega t + \phi)$, where m and n are positive integers. Consider displaying these motions on a single graph with x and y axes oriented perpendicular to each other. What restrictions are necessary on m, n, and ϕ if the trajectory is to be a closed curve? These curves are called *Lissajous figures*. Plot these curves for (*a*) $m = n = 1$, $\phi = 0$; (*b*) $m = n = 1$, $\phi = \pi/4$; and (*c*) $m = 2$, $n = 3$, $\phi = 0$. ••

82. What would be the period of oscillation of this book if you held it at one corner between index finger and thumb and allowed it to swing slightly? To first order, ignore any effect on the oscillation by friction between your fingers and the book. ••

83. You have a teardrop-shaped 2.00-kg object that has a hook at one end and is 0.28 m long along its longest axis (Figure P15.83). When you try to balance the object on your fingers, it balances when

Figure P15.83

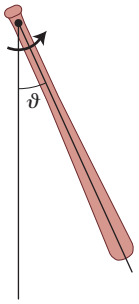


your fingers are 0.20 m from the hook end. When you hang the object by the hook and set it into simple harmonic motion, it completes 10 cycles in 11 s. What is its rotational inertia I about the hook? ••

84. A 2.00-kg object is free to slide on a horizontal surface. The object is attached to a spring of spring constant 200 N/m, and the other end of the spring is attached to a wall. The object is pulled in the direction away from the wall until the spring is stretched 50.0 mm from its relaxed position. The object is not released from rest, but is instead given an initial velocity of 2.00 m/s away from the wall. (*a*) Determine the energy of the oscillating system. (*b*) Write the equation for the object's position as a function of time. Ignore friction. ••

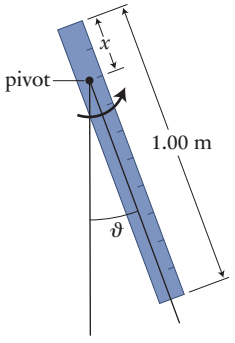
85. A hole is drilled through the narrow end of a 0.960-kg baseball bat, and the bat is hung on a nail so that it can swing freely (Figure P15.85). The bat is 0.860 m long, the hole is drilled 0.0300 m from the end, and the center of mass is located 0.670 m from the end. If the period of oscillation is 1.85 s, what is the bat's rotational inertia (*a*) about the pivot and (*b*) about the center of mass? (*c*) If the hole had been drilled 0.200 m from the end, what would the period be? ••

Figure P15.85



86. A meter stick is free to pivot around a position located a distance x below its top end, where $0 < x < 0.50$ m (Figure P15.86). (*a*) What is the frequency f of its oscillation if it moves as a pendulum? (*b*) To what position should you move the pivot if you want to minimize the period? ••

Figure P15.86



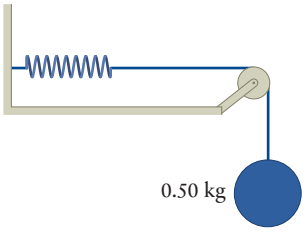
87. To determine what effect a spring's mass has on simple harmonic motion, consider a spring of mass m and relaxed length ℓ_{spring}. The spring is oriented horizontally, and one end is attached to a vertical surface. When the spring is stretched a distance x, the potential energy of the system is $\frac{1}{2}kx^2$. If the free end of the spring is moving with speed v when the free end is at x, calculate the kinetic energy of the spring in terms of m and v. (Hint: Divide the spring into many infinitesimal elements each of length $d\ell_{\text{segment}}$ and located at position ℓ_{segment}; then determine the speed of each piece in terms of ℓ_{segment}, $d\ell_{\text{segment}}$, m, v, and x and integrate from $\ell_{\text{segment}} = 0$ to $\ell_{\text{segment}} = \ell_{\text{spring}} + x$.) •••

88. (*a*) Show that for an object moving in simple harmonic motion, the speed of the object as a function of position is given by $v(x) = \omega\sqrt{A^2 - x^2}$. (*b*) Noting that $v_x = dx/dt$, isolate dt and integrate to determine how long it takes for an oscillator to move from its equilibrium position to some arbitrary position $x < A$. Compare this result with the value of $\Delta t = t_{x<A} - t_{x=A}$ calculated from Eq. 15.6, $x(t) = A\sin(\omega t + \phi_i)$. •••

89. The King's clock fails to keep time, and because you designed it, he's holding you responsible. Examining the spring-ball system you used as the central timekeeping oscillator, you realize that a simple pendulum might be more reliable, and there happens to be a bit of thin, strong wire handy. Unfortunately, you don't have any way of timing the oscillations, nor can you remember what the period is supposed to be. You don't even have a ruler to measure the wire for the pendulum, and there is not enough wire on hand to rely on trial and error. You suddenly realize that all you need to do is observe how far the ball stretches the spring when the spring is held vertically! You calm the King's wrath by putting your new design for the pendulum timekeeping mechanism in operation. ••• CR

90. In physics lab, you are measuring the period of a vertically hung spring-ball system in which the mass of the ball is 0.50 kg. As the system oscillates up and down, it also swings from side to side, a motion that interferes with your measurements of the vertical motion. After running several trials and calculating the period for the system set up this way, you decide to eliminate the side-to-side motion by rigging up the spring horizontally with a string draped over a pulley of radius 0.035 m, as in Figure P15.90. With this arrangement, your measured period is 10% greater than the value you measured for the vertically hung system. Realizing an opportunity for extra credit, you wonder if the rotational inertia of the pulley is buried in your data. ••• CR

Figure P15.90



91. You are aboard the International Space Station and are required to keep track of your mass because of your long stay in the free-fall conditions of deep space. To carry out this task, you hook a 215-kg cargo box to one end of a stiff spring and hook yourself to the spring's other end. With the spring and box initially at rest relative to you, you push off from the cargo box. Floating inside the cabin, you and the cargo box oscillate with a period of 2.21 s. You then unhook yourself from the spring and attach in your place a cargo box identical to the one at the other end of the spring. The oscillation period for this two-box system is 3.13 s. Amazingly, even with squeeze-tube food, you seem to be gaining mass. ••• CR



16

Waves in One Dimension

16.1 Representing waves graphically
16.2 Wave propagation
16.3 Superposition of waves
16.4 Boundary effects

16.5 Wave functions
16.6 Standing waves
16.7 Wave speed
16.8 Energy transport in waves
16.9 The wave equation

CONCEPTS

QUANTITATIVE TOOLS

Many phenomena around us involve waves. Throw a stone into a pond and ripples move outward along the surface of the water. Clap your hands and a sound wave travels outward. A large earthquake is detected all over the world because the earthquake causes a wave to propagate through Earth's interior. In general, a **wave** is a *disturbance* propagating (moving) either through a material or through empty space.

The most important feature of waves is that they do not carry matter from one place to another. After a wave has passed through a medium, the medium looks just as it did before it was disturbed. You prove this every time you and a stadium full of other fans work together to create a wave (**Figure 16.1**). Once you've done your small part—stand up, sit down—you are sitting in the same seat; in other words, you haven't moved from one location to another in the stadium. Or throw a rock into a still pond and watch the motion of a floating leaf: The leaf rises and falls as the wave created by the rock passes by but never moves away from its original position. Once the pond is still again, the leaf is just where it was before. A little thought will tell you that the same is true for every molecule of water in the pond. As we shall see, momentum and energy—not matter—are what move along with any wave.

Waves are collective phenomena: A stadium full of fans can create a wave that travels around the stadium; a single person, on the other hand, cannot create such a wave. The person can move up and down or run from one place to another, but neither motion is a wave. A wave can thus never exist at a single position in space: A single water molecule does not make a wave—a water wave requires an extended volume of water.

The examples with which I opened this chapter—water waves, sound waves, seismic waves—are all *mechanical waves*, which are waves that involve the (temporary) displacement of particles. Such waves require a *medium*—a substance to propagate along or through. The propagation is caused by the interaction of the particles in the medium. The medium for water ripples is the water; for the sound of clapping hands, the medium is the air; for seismic waves, it is the material that makes up Earth's interior. Some waves, such as radio waves and light, do not require any medium; they can propagate through empty space as well as through a medium. Even though such *nonmechanical waves* are very different in origin and nature from the mechanical waves we discuss in this chapter and the next, they share many of the properties of mechanical waves. For this reason, when we deal with nonmechanical waves in later chapters, we shall carry over many results derived for mechanical waves.

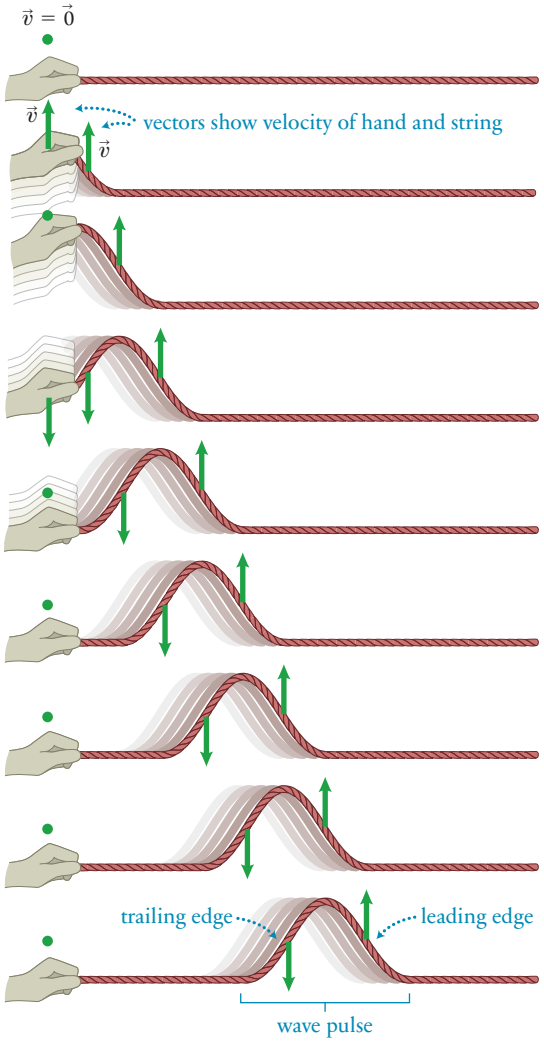
Figure 16.1 Fans at a stadium creating a wave. People get up and raise their arms as soon as the person next to them does so. Because of the small delay between one person and the next, not everyone gets up at the same instant, and the wave travels around the stadium at some finite speed.



16.1 Representing waves graphically

Suppose one end of a long, taut string is rapidly displaced up and down once, as shown in **Figure 16.2**. As the end of the string is raised, it pulls up the neighboring

Figure 16.2 Generation of a wave pulse that propagates along a string.



CONCEPTS

part of the string. As that neighboring part begins moving up, it pulls up the next part of the string, which, in turn, pulls up the next part, and so forth. The upward displacement of the end of the string is thus transmitted along the string because neighboring pieces of the string pull on each other. As the end of the string is moved back down, a similar chain of events occurs in the reverse direction. This pulling along of neighboring pieces of string, first up and then down, produces a **wave pulse**, which travels along the string at constant speed without changing shape. It is important that you do not confuse the horizontal motion of the pulse with the vertical-only motion of the string. The medium through which the pulse passes does not move in the horizontal direction. Each particle of the string moves only once up and once down as the pulse passes through it. **Figure 16.3** shows this vertical-only motion for a string of beads. (Although the string moves in two dimensions, up and down, the wave itself propagates in only one dimension.)

16.1 (*a*) Using the choice of axes shown in Figure 16.3, draw a position-versus-time graph showing how the x and y components of the position of the large bead change with time. (*b*) Are the x and y components of the velocity of the large bead positive, zero, or negative while the bead is on the leading edge of the pulse? Repeat for the trailing edge.

Figure 16.3 Video frames sequence of a wave pulse propagating along a string of beads.

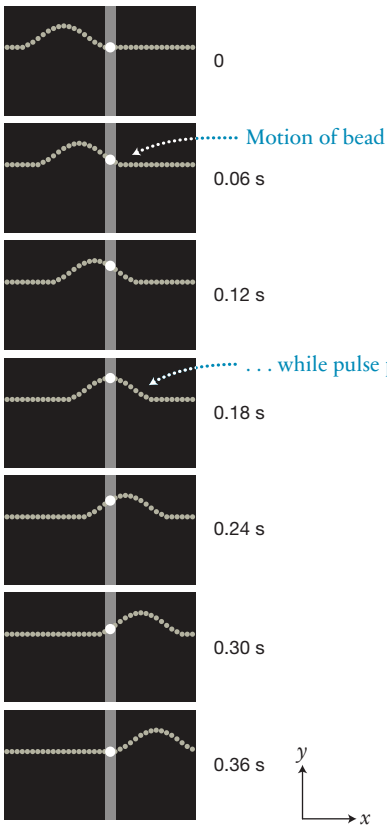


Figure 16.3 shows that the wave motion is different from the motion of the particles of the medium through which the pulse travels. The beads move in the vertical direction only, with each bead executing the same motion—first up and then down. Each bead moves at a different instant, however: first bead 1, then bead 2, then bead 3, and so on. At the leading edge of the pulse, each bead's motion is upward; at the trailing edge, it is downward. After the pulse has passed, all the beads are back in their initial positions.

An important point to keep in mind therefore is:

The motion of a wave (or of a single wave pulse) is distinct from the motion of the particles in the medium that transmits the wave (or pulse).

To emphasize this point, we shall use the letter c to denote the **wave speed** of a wave through a medium (along a string in this case) and v to denote the speed of the particles in the medium (the vertical motion of the beads in this case).*

16.2 (*a*) Is the wave speed c of the pulse in Figure 16.3 constant? (*b*) Determine the wave speed if the distance between the centers of adjacent beads is 5.0 mm.

As this checkpoint shows:

The wave speed c of a wave pulse along a string is constant.

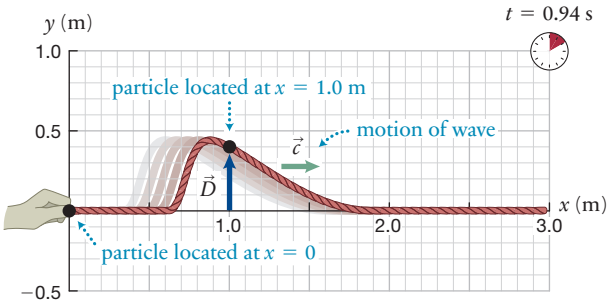
The motion of the various parts of the string, however, is not so simple. As the pulse propagates down the string, the displacements of various particles depend on time *and* on position along the string. At any given instant, different particles on the string have different displacements. Figures 16.2 and 16.3 show "snapshots" of a wave pulse along a string. Each snapshot shows the displacements of the particles on the string from their equilibrium positions as a function of position at one instant. For example, **Figure 16.4*a*** on the next page shows a snapshot of a triangular wave pulse traveling along a string. Such a triangular wave pulse is created by moving the end of the string first upward and then more rapidly downward. The snapshot is superimposed on a calibrated grid so that we can read off the displacements of various particles of the string from their equilibrium positions. The vector $\vec{D}$ in Figure 16.4*a* represents the **displacement** of the particle at $x = 1.0$ m at instant $t = 0.94$ s. In the context of waves, displacement is denoted by the symbol $\vec{D}$, and the initial position is always taken to be the equilibrium position of the particle.

*We use c for the speed of *any* wave, not just light (which is also a wave). The important point is never to confuse the speed of a wave with the speed of the particles being disturbed by the wave.

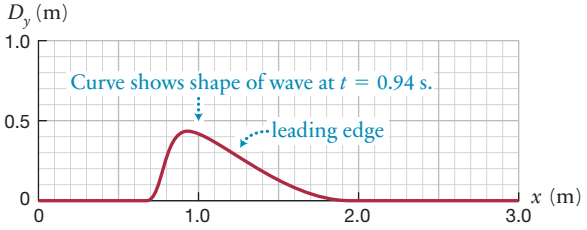
CONCEPTS

Figure 16.4 Distinction between the wave function and displacement curves for a triangular wave pulse propagating along a string parallel to the x axis.

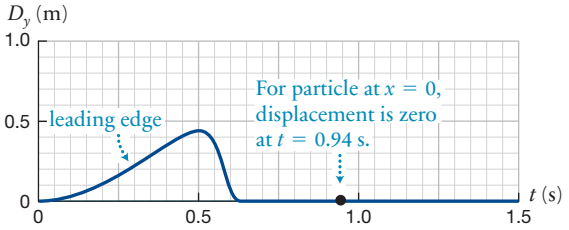
(a) Snapshot of wave at $t = 0.94$ s



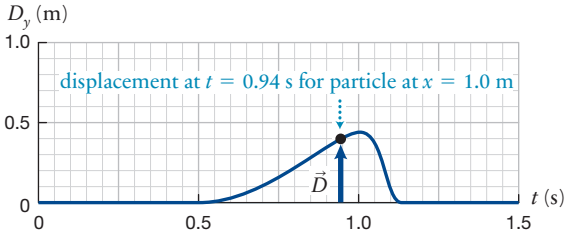
(b) Wave function at $t = 0.94$ s



(c) Displacement curve for particle located at $x = 0$



(d) Displacement curve for particle located at $x = 1.0$ m



The graphic representation of all the particle displacements at this instant is shown in Figure 16.4b, which gives the y components of the displacements of the particles of the string as a function of position x along the string (the left end of the string being at $x = 0$). The function shown in Figure 16.4b is called the **wave function**. The wave function represents the shape of the wave at a given instant and changes with time as the wave travels along the string.

16.3 From Figure 16.4a, determine the displacement of a point (a) 1.0 m, (b) 1.5 m, and (c) 2.0 m from the left end of the string. (d) On a snapshot taken a short time interval later, is the displacement at $x = 1.5$ m greater than, equal to, or smaller than that shown in Figure 16.4a?

A complementary way of representing the wave pulse of Figure 16.4a is to plot the displacement of one particle on the string as a function of time. Figure 16.4c, for example, shows the y component of the displacement of the end of the string at $x = 0$ as a function of time. The function shown in Figure 16.4c is called the **displacement curve.** The graph shows that the end of the string started moving up at $t = 0$, reached a maximum displacement of about $+0.50$ m at about $t = 0.50$ s, and then quickly moved back down to its initial position, which it reached at about 0.63 s. This is the motion that generated the triangular wave pulse propagating along the string.

The displacement curve for any string through which a wave pulse moves is different for different particles on the string, as you can see by comparing Figure 16.4c and d. Figure 16.4d shows the displacement curve for the particle at $x = 1.0$ m. The displacement remains zero until the leading edge of the pulse has traveled from $x = 0$ to $x = 1.0$ m. This happens at $t = 0.50$ s, and at that instant the leading edge of the pulse begins raising the $x = 1.0$ m particle until it reaches its maximum displacement at $t = 1.0$ s. The much steeper trailing end of the pulse shown in Figure 16.4a then quickly returns the particle to its equilibrium position between $t = 1.0$ s and $t = 1.1$ s.

Why is the displacement curve the mirror image of the wave function? Because the wave pulse travels to the right, the leading edge of the pulse is on the right on the wave function graph. The displacement curve, on the other hand, has small values of t (earlier instants) on the left and large values of t (later instants) on the right. So in Figure 16.4c the leading edge of the pulse is on the *left*. (To convince yourself, look at the particle at $x = 1.0$ m in Figure 16.4a and visualize the motion of that particle as the pulse travels past it. Sketch the displacement of the particle as a function of time and compare your result with Figure 16.4d.)

Note that the displacement curves in Figures 16.4c and d have the same form: The one for $x = 1.0$ m is simply shifted to the right. This tells us that both particles execute the same motion at different instants.

16.4 (a) From Figures 16.4c and d, determine how long it takes the pulse to travel from $x = 0$ to $x = 1.0$ m. (b) Using your answer to part a, determine the wave speed.

CONCEPTS

Figure 16.5 Longitudinal wave pulse propagating along a spring.

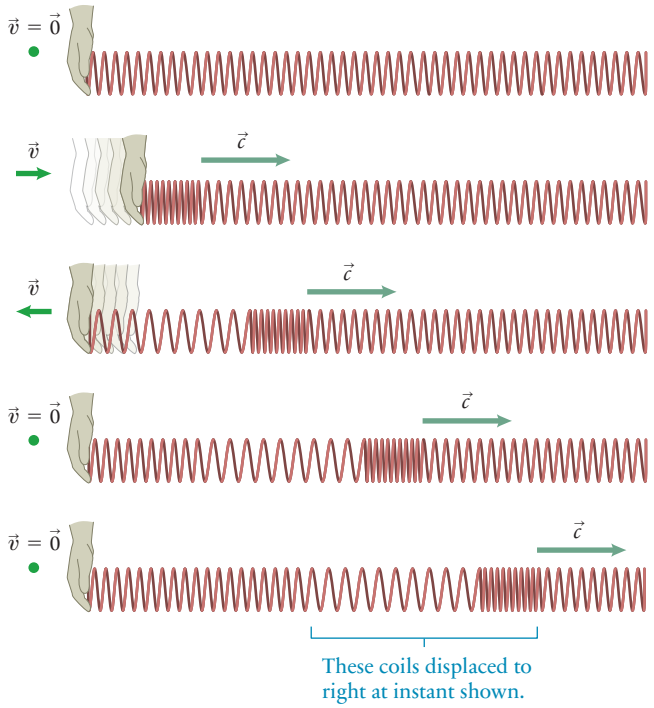
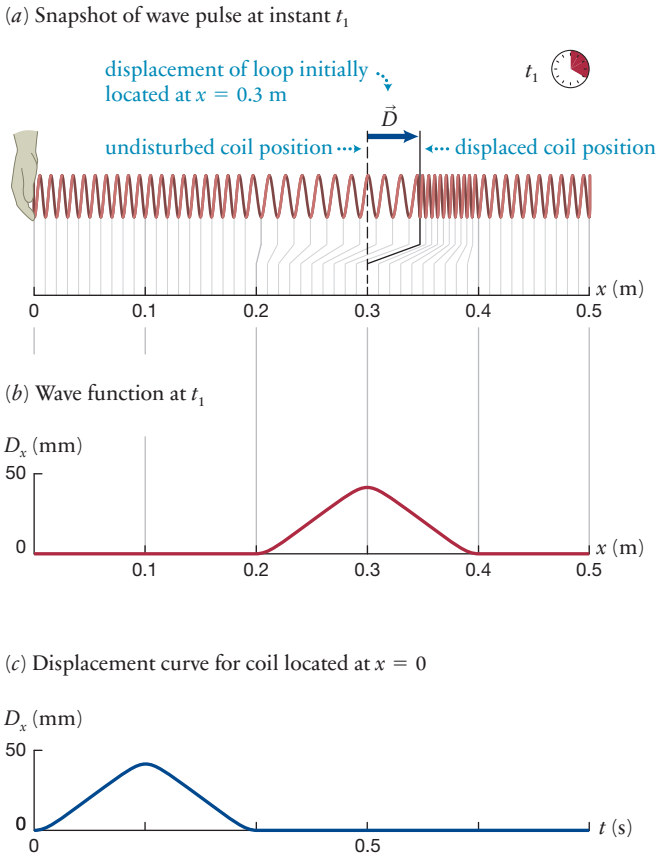


Figure 16.6 Wave function and displacement curve for a longitudinal wave pulse propagating along a spring.



Mechanical waves are divided into two categories, depending on how the medium moves relative to the wave motion. For a wave propagating along a string, the medium movement is perpendicular to the pulse movement. Such waves are called **transverse waves**. In the other type of mechanical wave, the medium movement is parallel to the pulse movement, and waves that behave this way are called **longitudinal waves**. Let's now look at this second type of wave in more detail.

Consider a long coil spring whose end is rapidly displaced back and forth as in **Figure 16.5**. As the left end of the spring is displaced to the right, each coil pushes on the next, displacing the coils one after the other to the right. When the left end is pulled back to its original position, each coil pulls the next one back. The result is again a wave pulse propagating to the right, with each coil temporarily being displaced in the direction of propagation of the wave pulse.

Even though the direction of motion of the individual coils coincides with the direction of propagation of the wave pulse, the motion of the coils and that of the pulse are still distinct. The individual coils move back and forth, but the wave pulse travels at constant speed along the spring.

Longitudinal waves can also be represented using wave functions and displacement curves, but now the wave functions no longer look like a snapshot of the propagating wave. **Figure 16.6*a*** shows a snapshot of a pulse propagating along a spring at some instant t_1, overlaid on a calibrated grid. The corresponding wave function is shown in Figure 16.6*b*—the wave function gives the x component of the displacement of each coil as a function of the coil's position. The coils between $x = 0$ and $x = +0.20$ m are undisturbed, and hence the displacement $\vec{D}$ is zero between $x = 0$ and $x = +0.20$ m. The coils between $x = +0.20$ m and $x = +0.40$ m are all displaced rightward. The x component of the displacement is positive because the displacement is in the positive x direction. To the right of $x = +0.40$ m, the displacement is again zero because the pulse has not yet reached that part of the spring.

Figure 16.6*c* shows the displacement curve for the end of the spring at $x = 0$.

16.5 (*a*) The pulse in Figure 16.6*b* is symmetrical even though Figure 16.6*a* shows that the length of the stretched portion of the spring is 0.15 m and the length of the compressed portion of the spring is only 0.05 m. Explain how the symmetrical curve of Figure 16.6*b* is a correct representation of the asymmetrical situation shown in Figure 16.6*a*. (*b*) Suppose that instead of the end of the spring moving to the right, as illustrated in Figure 16.6, the end is moved to the left. Would this disturbance still cause a wave pulse to travel to the right? If yes, sketch the wave function for this wave pulse at some instant $t > 0$. If no, what happens instead?

CONCEPTS

16.2 Wave propagation

A wave pulse is not an object—it has no mass—and so the description of wave motion is very different from the description of the motion of objects. To study the propagation of a wave pulse along a string, let us consider a collection of beads connected by short strings. The mass of the strings is negligible relative to that of the beads. **Figure 16.7*a*** shows the first ten of these beads as a pulse travels along the string. The vertical gray lines show that the motion of the beads is entirely vertical.

Let us focus on the third bead along the string. At t_1 the left end of the string has just begun moving up, and only bead 1 is displaced from its initial position. The top free-body diagram in Figure 16.7*b* shows the two contact forces exerted on bead 3. If the string is held taut, $\vec{F}^{\text{c}}_{23}$ and $\vec{F}^{\text{c}}_{43}$ point in opposite horizontal directions. (This amounts to saying that the force of gravity on the beads is negligible relative to the contact forces exerted by the strings, and because the bead does not accelerate, the vector sum of the forces exerted on it must be zero.)

16.6 Answer these five questions about the situation at t_1 in Figure 16.7, assuming the right end of the string of beads is attached to a wall that is not shown. (*a*) If $F^{\text{c}}_{43} = 5$ N, what is the magnitude of the force exerted by the wall on the rightmost bead? (*b*) What is the magnitude of the horizontal component of $\vec{F}^{\text{c}}_{12}$? (*c*) What is the direction of the force exerted by the hand on the string? (*d*) How does the magnitude of this force compare with the magnitude F^{c}_{43}? (*e*) What are the effects of the horizontal and vertical components of this force?

This checkpoint makes an important point: If the string is held taut, the magnitudes of the horizontal components of all the contact forces between adjacent beads are equal to the tension $\mathcal{T}$ in the string.

At t_2 the leading edge of the pulse has propagated to bead 3. Because of the displacement of bead 2, $\vec{F}^{\text{c}}_{23}$ is no longer horizontal. The displacement of 2 also causes the string between 2 and 3 to stretch, and so the magnitude F^{c}_{23} increases.

The magnitude of the horizontal component of $\vec{F}^{\text{c}}_{23}$ is still equal to $\mathcal{T}$, but the vector sum of the forces exerted on bead 3 is now nonzero and points vertically upward. Consequently, bead 3 now has a nonzero upward acceleration and so begins moving upward (Figure 16.7*c*). As it moves upward, it stretches and changes the direction of the string connecting it to bead 4, and so the magnitude and direction of $\vec{F}^{\text{c}}_{43}$ change. These changes in $\vec{F}^{\text{c}}_{43}$ have two effects: $\vec{F}^{\text{c}}_{43}$ begins to slow down bead 3, and the reciprocal force $\vec{F}^{\text{c}}_{34}$ begins to accelerate bead 4 upward.

At some instant between t_3 and t_4, $\vec{F}^{\text{c}}_{23}$ and $\vec{F}^{\text{c}}_{43}$ again become equal in magnitude and opposite in direction (but this time with $\vec{F}^{\text{c}}_{23}$ directed upward and to the left and $\vec{F}^{\text{c}}_{43}$ directed downward and to the right), and so the vector sum of the forces exerted on bead 3 is zero. As the bead continues to move upward, the vector sum of the forces begins to oppose the bead's motion (free-body diagram at t_4). Consequently the bead slows down until it comes to a stop at t_5. At that instant, the bead has reached its maximum displacement, and the vector sum of the forces exerted on it points downward. As the bead moves back down, the vector sum of the forces exerted on it decreases. At t_7 the vector sum points upward again, slowing the downward motion of the bead until it reaches its initial position again at t_8. Once the pulse has moved beyond bead 3 and the string around it is straight again, the vector sum of $\vec{F}^{\text{c}}_{23}$ and $\vec{F}^{\text{c}}_{43}$ is again zero, and so the bead remains at rest.

16.7 (*a*) Note that the displacement of bead 4 at t_4 in Figure 16.7*a* is the same as that of bead 3 at t_3. How do the velocity and acceleration of bead 4 at t_4 compare with those of bead 3 at t_3? (*b*) You move one end of two different strings, A and B, up and down in the same way. Suppose the resulting pulse travels twice as fast on string A as on string B. How does the velocity of a particle of string A compare with the velocity of a particle of string B that has the same nonzero displacement? (*c*) Sketch the pulses as they propagate along the two strings and point out any differences between the two. (*d*) For each string, consider a particle located 0.20 m from the left end. Sketch a displacement-versus-time graph for each particle, and point out any differences between your two graphs.

Be sure not to skip Checkpoint 16.7 because it makes two very important—and somewhat counter-intuitive—points:

When a particle of the string is displaced from its equilibrium position, its velocity $\vec{v}$ and acceleration $\vec{a}$ are determined only by the initial disturbance and are independent of the wave speed *c*.

All particles of the string execute the same motion as that caused by the initial disturbance, *regardless of wave speed c* (which determines only how quickly they execute this motion after one another). As Checkpoint 16.7 shows, however, the *shape* of the pulse depends on *c*:

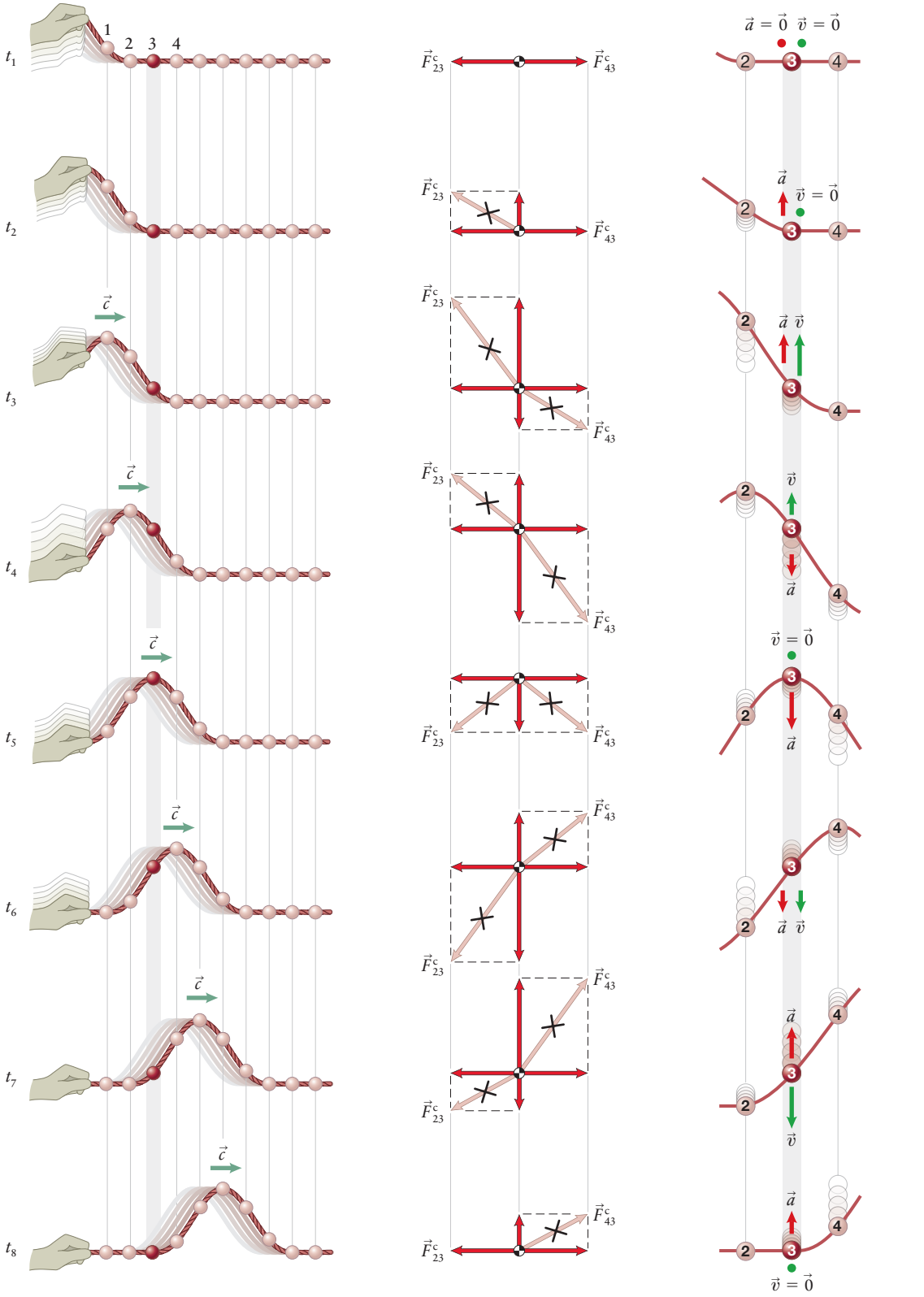
For a given disturbance, high wave speeds yield wave pulses that are stretched out and low wave speeds result in pulses that are more compressed.

Regardless of the pulse shape, however, any particle of the string that reaches a certain displacement has the same velocity and acceleration as any other particle at the instant it reaches that same displacement in the same part of the pulse.

What determines the wave speed? Our analysis of Figure 16.7 shows that the propagation of a wave

CONCEPTS

Figure 16.7 Forces and accelerations that cause a wave pulse to propagate along a string of beads.



pulse along a string is due to the forces exerted by each piece on the neighboring pieces. This suggests that wave speed is related to the strength of these forces. Let us therefore investigate what happens if we change the magnitude of the tension in the string.

Suppose that at instant t_3 in Figure 16.7, when bead 3 is on the leading edge of the pulse, this bead has an acceleration $\vec{a}$ when its vertical displacement is $\vec{D}$ (**Figure 16.8a**). This acceleration is caused by the vector sum $\Sigma\vec{F}$ of the forces exerted on the bead. If we now increase the tension in the string (Figure 16.8*b*), the magnitudes of the horizontal components of the contact forces increase by the same amount. Because the contact forces are now greater, the vector sum of the forces exerted on bead 3 reaches the same value as in Figure 16.8*a* with a smaller displacement of bead 2. In other words, it takes a shorter time interval for the pulse to propagate from one bead to the next. A greater tension in the string thus causes a higher wave speed.

Increasing the mass of each bead accomplishes the opposite effect: For a given tension, increasing the mass of each bead decreases the beads' accelerations. To give this larger bead 3 the same acceleration $\vec{a}$ that the smaller bead 3 in Figure 16.8*a* has, bead 2 must now be displaced farther. Thus now it takes a longer time interval for the pulse to propagate from one bead to the next (Figure 16.4*c*). Increasing the mass of each bead corresponds to increasing the mass per unit length of a continuous string.

The speed *c* of a wave propagating along a string increases with increasing tension in the string and decreases with increasing mass per unit length along the string.

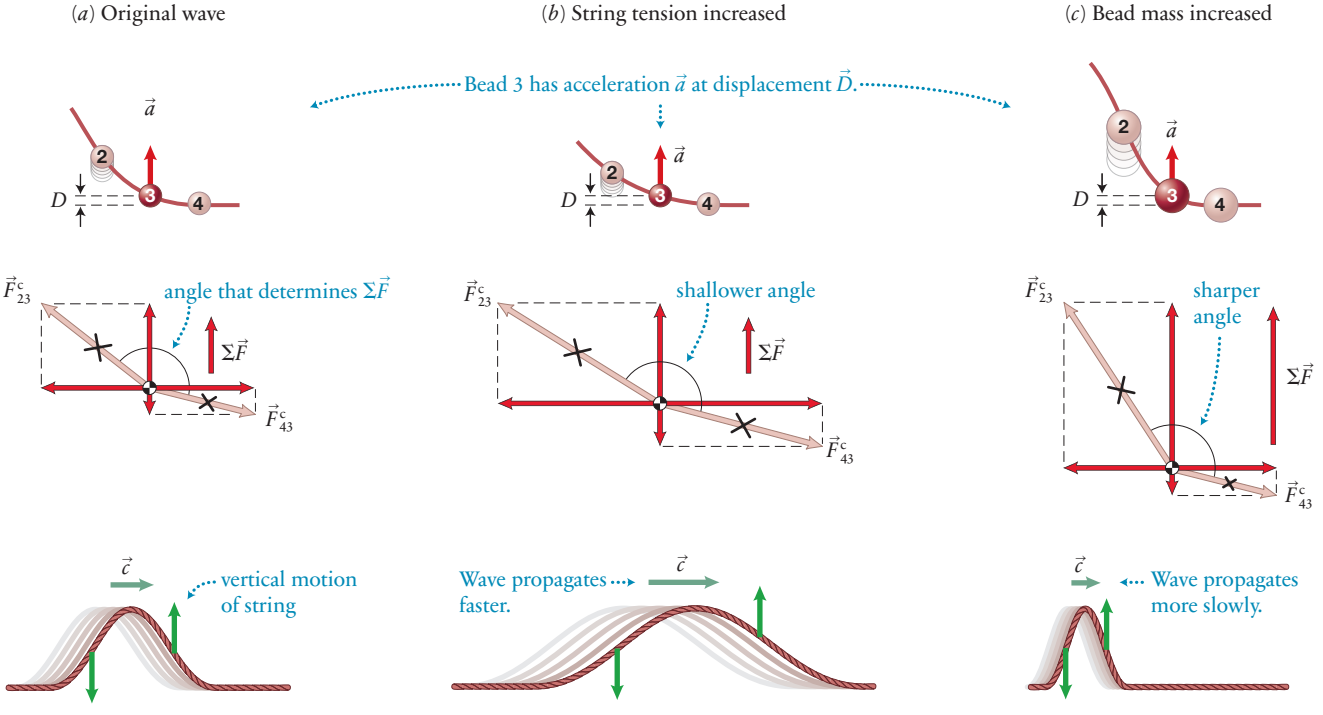
This result makes sense: A greater tension means a greater restoring force and therefore a greater tendency to pass the displacement along, resulting in a higher wave speed. If the mass is greater, there is a greater resistance to passing the displacement along, and so the wave speed is lower.

16.8 Two strings, A and B, are identical except that the tension in A is greater than that in B. Suppose you move the left end of each string rapidly up and down once. For each string, sketch (*a*) the wave functions at the instant the pulse on A has traveled halfway down the length of the string and (*b*) the displacement curve for a particle midway down the string.

If the end of a string is made to execute a periodic motion, the resulting wave is called a **periodic wave.** **Figure 16.9** shows a special type of periodic wave, called a **harmonic wave,** obtained by moving the end of the string so that it oscillates harmonically. The result is a sinusoidally shaped wave traveling along the string.

16.9 Plot the displacement of the left end of the string in Figure 16.9 as a function of time. What similarities and differences exist between your graph and the shape of the wave in Figure 16.9?

Figure 16.8 Effects of tension and mass on wave speed. (*a*) The free-body diagrams show the angle between $\vec{F}^c_{23}$ and $\vec{F}^c_{43}$ that is required to cause an acceleration $\vec{a}$ at a displacement $\vec{D}$. (*b*) When the tension is greater, a shallower angle is required to produce the same condition so the wave propagates faster. (*c*) When the mass is greater, a sharper angle is required to produce the same condition so the wave propagates more slowly.



CONCEPTS

Note how the shape of the string in Figure 16.9 resembles the curves representing simple harmonic oscillations from Chapter 15. The shape in Figure 16.9, however, represents the displacement of the medium as a function of *position*, whereas the simple harmonic oscillator curves in Chapter 15 represented displacement as a function of *time*. If you focus on one particle of the string as the wave moves through it, however, you would see that its up-and-down motion is a delayed version of the hand's periodic motion. So a number of quantities we have used to describe periodic motion are also useful to describe waves (*amplitude:* maximum displacement of any given particle of the medium from its equilibrium position; *frequency:* number of cycles per second executed by each particle of the medium; *period:* time interval taken to complete one cycle). A periodic wave repeats itself over a distance called the **wavelength**, denoted by the symbol λ (Figure 16.9). Each time the left end of the string executes a complete oscillation, the wave advances by one wavelength. If the speed at which the wave propagates is c, then the wave advances a distance cT during the period T of the sinusoidal motion. Therefore:

The wavelength of a periodic wave is equal to the product of the wave speed and the period of the wave motion.

Contrary to what you may expect, moving your hand up and down more quickly does not generate a faster-traveling pulse. **Figure 16.10** shows two pulses generated on identical strings held under identical tensions. The end of the string in Figure 16.10*b* is displaced up and down more quickly than that in Figure 16.10*a*. Both pulses have the same displacement $\Delta\vec{r}$ during a time interval Δt, so the wave speed c is the same for both pulses. To a very good approximation, we determine experimentally that:

The speed c of a wave propagating along a string is independent of the velocities $\vec{v}$ of the individual pieces of string. The value of c is determined entirely by the properties of the medium.

Figure 16.9 If the end of a string is made to execute a simple harmonic motion, the resulting traveling wave is sinusoidal in shape. During one period of the motion, the wave advances by one wavelength.

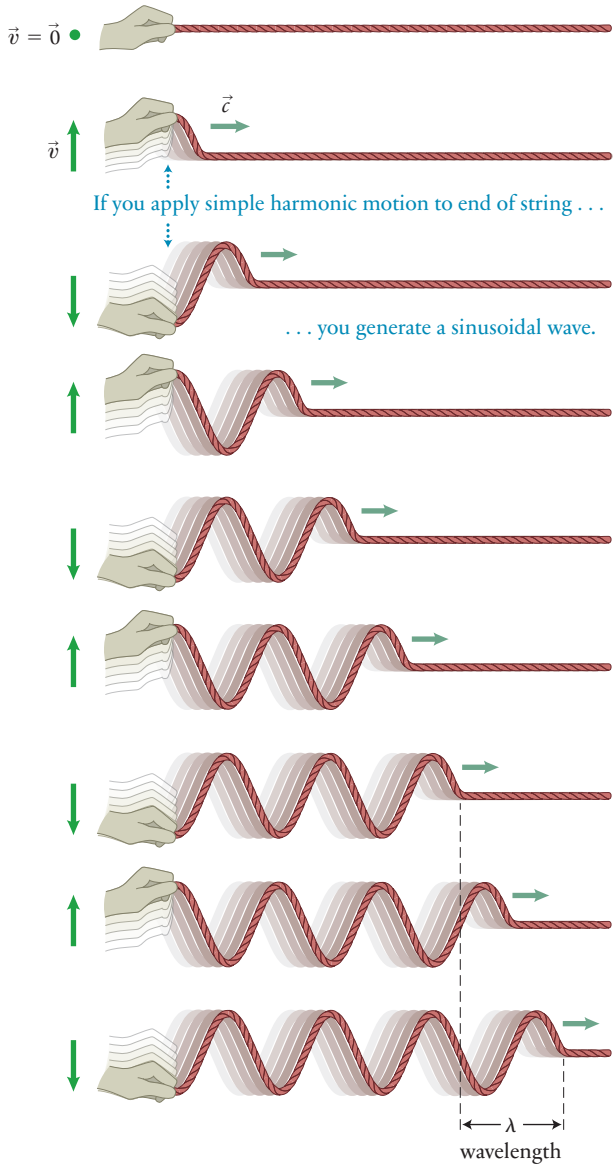
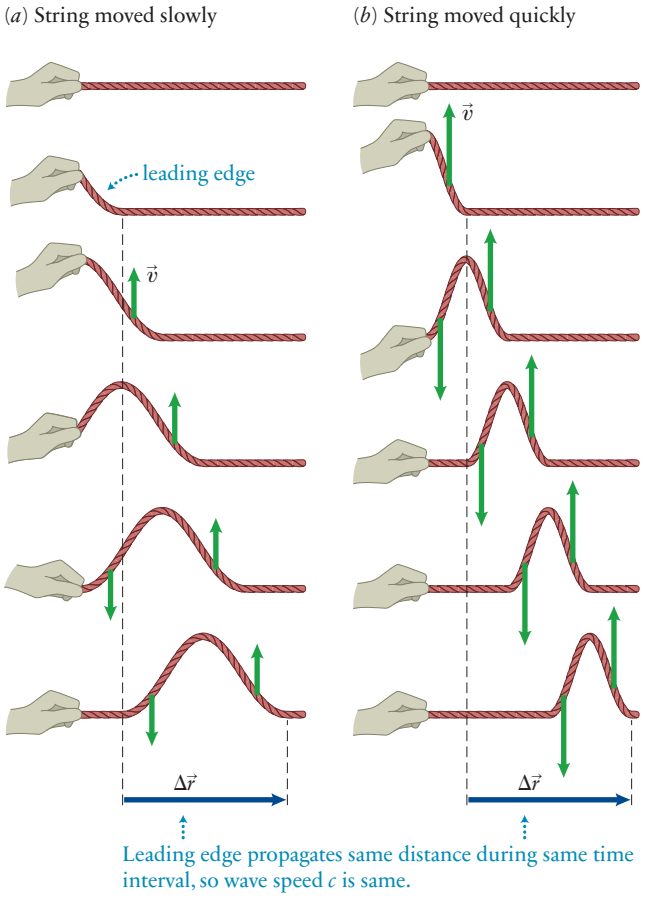


Figure 16.10 The effect on a wave pulse of the speed with which the hand moves up and down. In each case the leading edge of the pulse propagates the same distance during the same time interval, showing that the wave speed c is independent of the velocities $\vec{v}$ of the particles in the string.



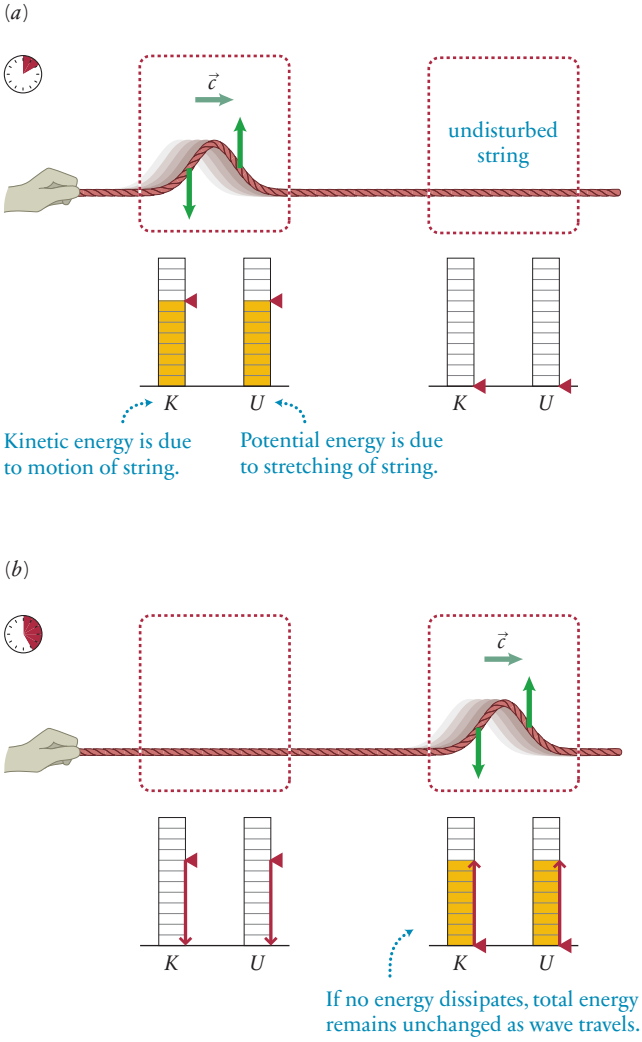
CONCEPTS

If wave speed depended on the velocity of the particles, the shape of a wave pulse would change as it propagates because some parts of it would propagate more quickly than others. As long as the displacement caused by the wave is not too large, however, the shape of the waves remains unchanged (see Figure 16.3).

16.10 Consider a harmonic wave traveling along a string. Are the following quantities determined by the source of the wave, by the properties of the string, or by both: (*a*) the period of oscillation of a particle of the string; (*b*) the speed c at which the wave travels along the string; (*c*) the wavelength; (*d*) the maximum speed v of a particle of the string?

Figure 16.11 illustrates how a propagating wave pulse carries two forms of energy along with it: kinetic energy associated with the motion of individual particles and elastic potential energy associated with the stretching of the string as the pulse is passing through. In Section 16.3 we shall see that a mechanical wave always carries equal amounts of these two forms of energy.

Figure 16.11 A wave pulse carries kinetic and potential energy. If there is no energy dissipation, the amount of energy in (*b*) is the same as that in (*a*).



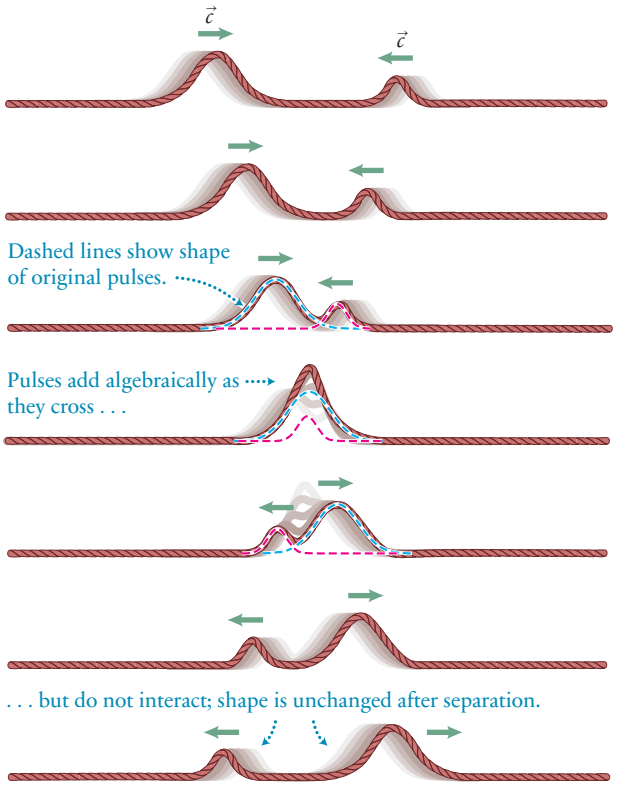
16.11 Does the wave pulse in Figure 16.11 also carry along momentum? Justify your answer.

16.3 Superposition of waves

Waves have a remarkable property: Two waves can pass straight through each other without changing each other's shape. **Figure 16.12** shows a series of snapshots of two pulses propagating toward each other along a string. When the two pulses overlap, the displacements caused by the two waves add algebraically, provided the medium obeys Hooke's law (see Section 8.9) for the combined displacement. This phenomenon is called the **superposition of waves:**

> **If two or more waves overlap in a medium that obeys Hooke's law, then the resulting wave function at any instant is the algebraic sum of the individual wave functions.**

Figure 16.12 Constructive interference. Two pulses propagating in opposite directions interfere constructively if they displace the string particles in the same direction (here, upward). When the pulses cross, their displacements add algebraically, giving rise to a displacement greater than that caused by either pulse individually. The pulses do not interact with each other: After they separate, their shapes are unaltered.



CONCEPTS

The resulting wave may have a very complicated shape when the waves overlap, but when they separate again, each wave still has its original shape.

The superposition of waves further distinguishes waves from objects: Two objects cannot occupy the same location in space at the same instant, but two waves *can* occupy the same location at the same instant.

Because the wave functions of two overlapping waves add algebraically, the resultant displacement is greater than that of either wave when the individual displacements have the same sign (Figure 16.12). When they have opposite signs (**Figure 16.13**), the wave with the smaller displacement decreases the displacement caused by the other wave.

When two waves overlap, we say that the waves *interfere* with each other, and the phenomenon is called **interference**. The adding of wave displacements of the same sign, as in Figure 16.12, is called *constructive interference*, and the adding of wave displacements of opposite signs, as in Figure 16.13, is called *destructive interference*. This terminology is somewhat misleading because the waves do not *interact* in any way. As you can see from Figures 16.12 and 16.13, when the two waves separate, their shapes are unchanged. Although there is no destruction or construction, interference of waves does affect the displacement that occurs at specific locations, as the next checkpoint shows.

16.12 (*a*) Is the maximum displacement the same for all particles of the string in Figure 16.12? For all particles of the string in Figure 16.13? (*b*) For each string, sketch a displacement curve for a point near the left end of the string. (*c*) Repeat part *b* for the point of each string at which the pulses meet. (*d*) Reconcile your answer to part *a* with your sketches for parts *b* and *c*.

An extreme form of destructive interference occurs when the two pulses crossing each other are the same size and shape but of opposite algebraic signs, as in **Figure 16.14**. At the instant the two pulses overlap, the displacements cancel and the string is flat. Even though it appears as though both pulses have disappeared, they continue unchanged after the superposition.

Figure 16.13 Destructive interference. Two pulses propagating in opposite directions interfere destructively if they displace the string particles in opposite directions. When the pulses cross, their displacements add algebraically, giving rise to a decreased displacement. After the pulses separate, their shapes are unaltered.

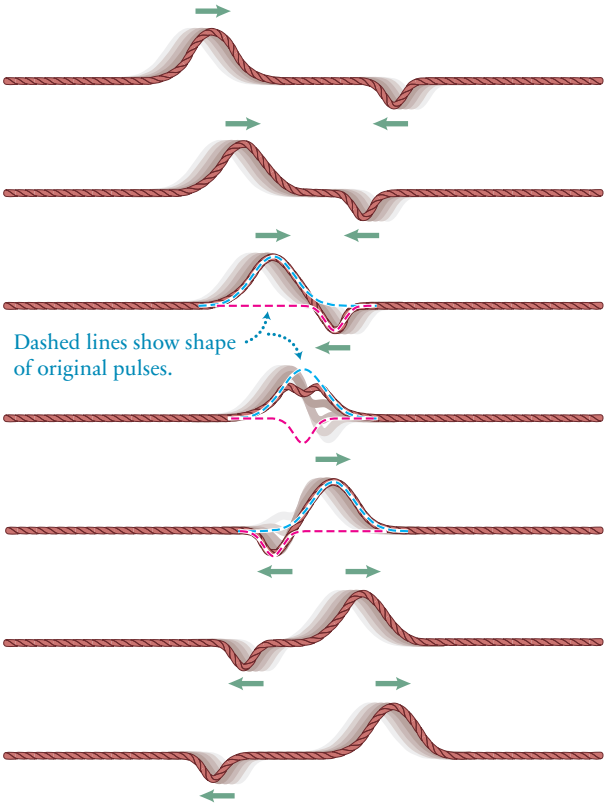
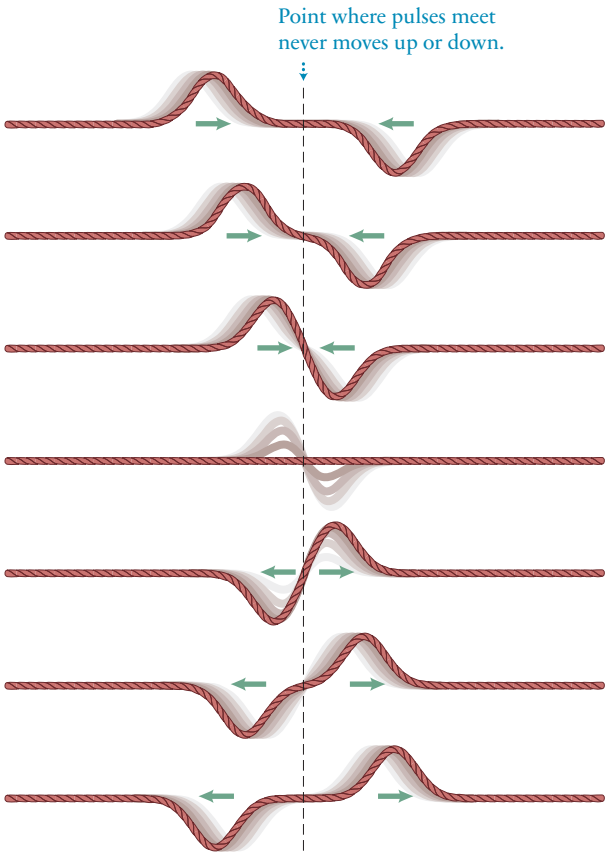


Figure 16.14 Complete destructive interference of two pulses of identical shape propagating in opposite directions while displacing string particles in opposite directions. At the instant the pulses overlap, their displacements add to zero.



CONCEPTS

16.13 (*a*) Sketch a displacement curve for the point on the string in Figure 16.14 at which the two pulses meet. (*b*) When two wave pulses overlap, how is the velocity of a point in the overlap region related to the velocities of the corresponding points of the individual pulses?

As Checkpoint 16.13 shows, the point at which the two pulses in Figure 16.14 meet never moves. When the two pulses overlap, they cause equal displacements in opposite directions, and so the vector sum of the displacements is zero. A point that remains stationary in a medium through which waves move is called a **node**.

We can deduce a useful fact by examining what happens to the energy in the pulses when they overlap exactly. Because the two pulses are identical, each carries the same amount of energy $E_1 = K_1 + U_1$, and the total energy in the two pulses is thus $2K_1 + 2U_1$ (**Figure 16.15a**). When the string displacements caused by the two pulses cancel each other exactly (the pulses interfere destructively), the string is straight, and so the elastic potential energy is zero. The kinetic energy, however, is not zero. The upward-moving leading edge of pulse 1 adds to the upward-moving trailing edge of pulse 2, and likewise for the downward-moving edges. The result is that the points within the overlapping pulses move twice as fast as those within the individual pulses before and after the overlap. Because kinetic energy depends on the square of the velocities, the kinetic energy of points in the overlap region is four times as great as that of an individual pulse. The energy during the instant shown in Figure 16.15*b* is thus $4K_1$, which must be equal to the initial energy $2K_1 + 2U_1$ of the two separate pulses. So we determine that $2K_1 + 2U_1 = 4K_1$, which can be true only if $K_1 = U_1$—in other words, if each pulse has an equal amount of kinetic and potential energy. Because these arguments do not depend on the shape of the pulse, it follows that:

A wave contains equal amounts of kinetic and potential energy.

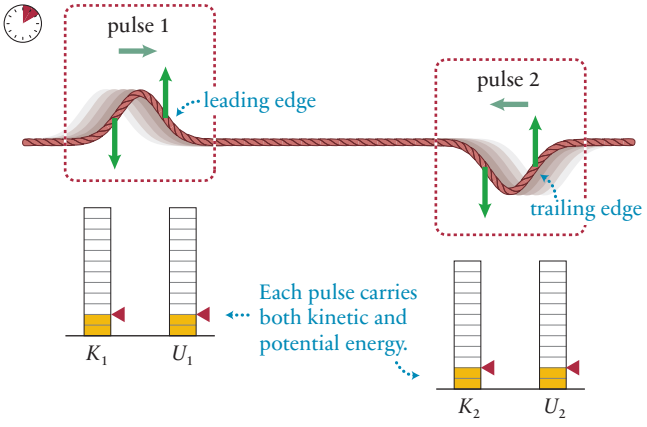
CONCEPTS

16.4 Boundary effects

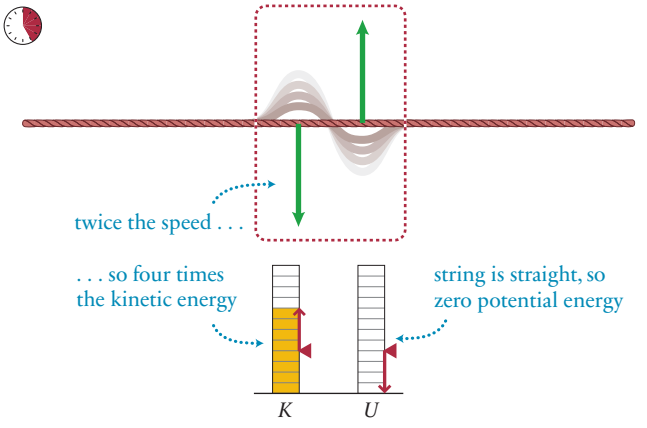
When a wave pulse reaches a boundary where the transmitting medium ends, the pulse is *reflected*, which means that its direction of propagation is reversed. After reaching the boundary, a pulse originally traveling to the right starts moving to the left through the medium, and a pulse originally traveling to the left starts moving to the right. Before the pulse reaches the boundary, we refer to it as the *incident pulse;* after reflection, it is the *reflected pulse*. Whether or not a reflected pulse has the same orientation as the incident pulse depends on the properties of the boundary.

Figure 16.15 Interference of two pulses traveling in opposite directions. At the instant the pulses overlap, the string displacement is zero everywhere, and so the potential energy is zero; all the energy is kinetic.

(*a*) Two pulses of identical shape traveling in opposite directions and having displacements in opposite directions



(*b*) At the instant the pulses overlap, the string displacement is zero everywhere, and so the potential energy is zero; all the energy is kinetic



Consider first a pulse propagating along a string toward the end of the string that has been anchored so that it cannot move (**Figure 16.16**). As the leading edge of the pulse reaches the fixed end, the pulse pulls the end upward. The effect of this upward pull is that the fixed end exerts a downward force on the string.* This downward force exerted on the string produces an *inverted* reflected pulse. If no energy is lost in the reflection, the reflected pulse is identical in shape to the incident pulse.

*If you think it is hard to imagine that the fixed end pulls the string downward, imagine pulling yourself up from a chinning bar. When you do so, you exert a downward force on the bar, and the bar exerts an upward force on you. Because the bar is fixed, you end up moving up even though you pull down on the bar.

Figure 16.16 Reflection of a pulse at a boundary where the string end is fixed and so cannot move.

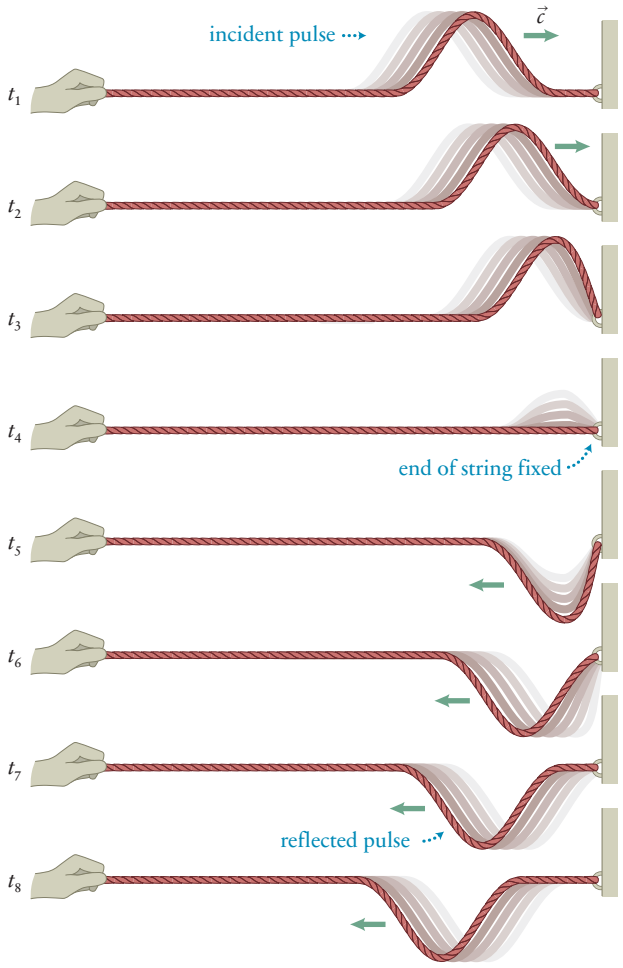
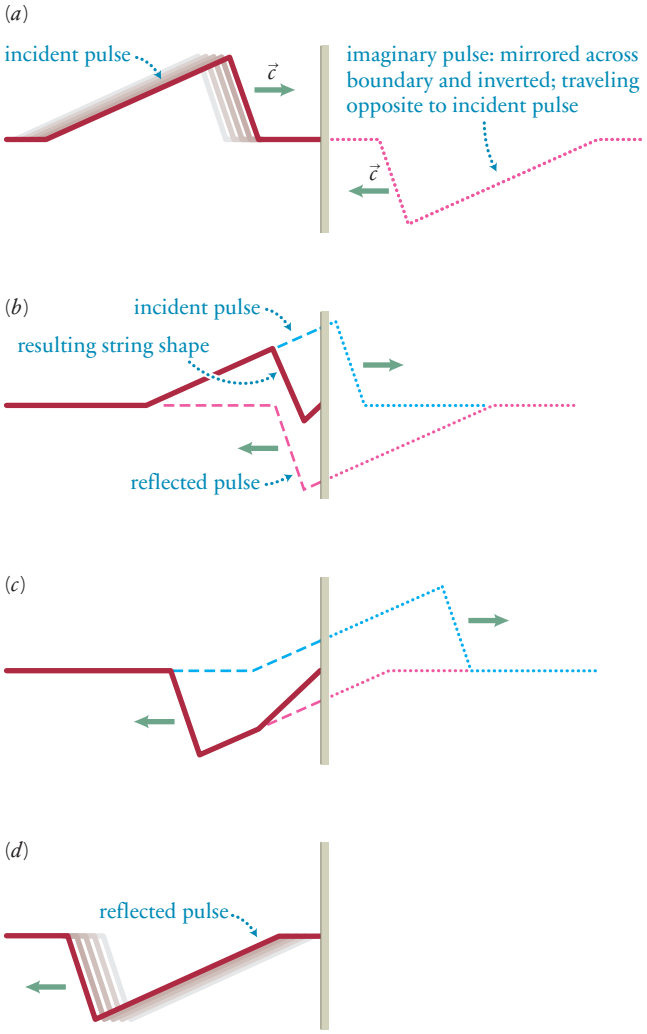


Figure 16.17 Reflection of a triangular pulse at a boundary where the string end cannot move.



16.14 The string in Figure 16.16 is perfectly straight at t_4. What has happened to the energy in the incident pulse at that instant?

Notice the similarity between Figures 16.16 and 16.14. If you cover the right half of Figure 16.14 with a piece of paper so that the left edge of the paper goes through the point at which the pulses meet, the two figures are identical. Indeed, as we saw in Checkpoint 16.13, the point at which the pulses meet in Figure 16.14 is a node, precisely like the fixed end of the string in Figure 16.16.

The correspondence between Figure 16.16 and the left side of Figure 16.14 suggests a procedure for determining the shape of a reflected wave pulse: On the incident pulse superpose an identical, but inverted, pulse that approaches from the opposite side of the fixed end and reaches the fixed end at the same instant as the incident pulse. This procedure is best illustrated using an example involving a pulse with an asymmetrical shape (**Figure 16.17**).

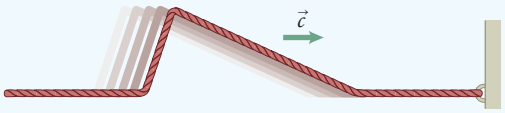
Figure 16.17*a* shows a triangular wave pulse approaching the fixed end of a string. To construct the reflected pulse we first draw an inverted triangular pulse on the opposite side of the fixed point and moving in the opposite direction. This inverted "reflected pulse" reaches the fixed point at the same instant as the incident pulse, and the two interfere (Figures 16.17*b* and *c*). While the two pulses overlap in space, the resulting shape of the pulse is obtained by adding the incident and reflected waves. Once the reflected pulse separates from the fixed end, its shape is the inverse of the incident wave pulse (Figure 16.17*d*).

CONCEPTS

EXAMPLE 16.1 Reflection from a fixed end

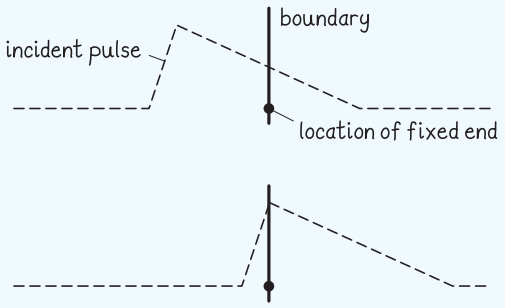
Consider a triangular wave pulse approaching the fixed end of a string (**Figure 16.18**). Sketch the shape of the string (*a*) when a point halfway up the leading edge of the pulse has reached the fixed end and (*b*) when the peak of the pulse has reached the fixed end.

Figure 16.18 Examples 16.1 and 16.2.



❶ **GETTING STARTED** I begin by sketching the pulse at the two instants given in the problem, pretending that the fixed end is not there and no reflection has occurred (**Figure 16.19**). I use a dashed line for the pulse to indicate that this is not the actual shape of the string at these instants (because reflection *has* occurred), and I draw a vertical line to indicate the (horizontal) position of the fixed end.

Figure 16.19



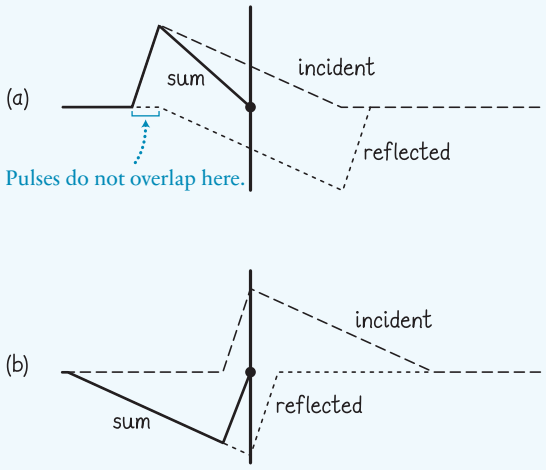
❷ **DEVISE PLAN** Once any part of the pulse reaches the fixed end, reflection begins. To determine the pulse shape during reflection, I must draw a reflected pulse that is both inverted and reversed relative to the incident pulse and then add that reflected pulse to my incident pulses in Figure 16.19.

❸ **EXECUTE PLAN** (*a*) At the fixed end, the sum of the displacement of the incident pulse and the displacement of the reflected pulse is zero. In the region where the two pulses do not overlap, the sum of the displacements of the two pulses is equal to the displacement of the trailing edge of the incident pulse. Between this no-overlap region and the fixed end, I add the dashed and dotted lines algebraically to determine the shape of the string. The resulting shape is shown in **Figure 16.20***a*. ✔

(*b*) Following the same procedure yields Figure 16.20*b*. ✔

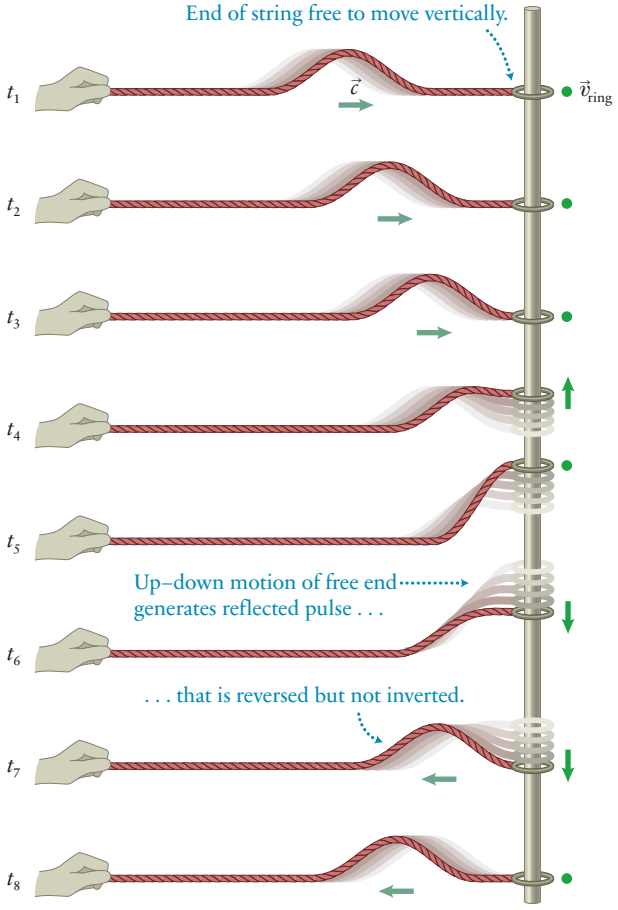
❹ **EVALUATE RESULT** The pulses I drew terminate at the node at the fixed end, as expected. In Figure 16.20*b* the string shape resembles the reflected pulse more than the incident pulse because most of the incident pulse has traveled past the fixed end.

Figure 16.20



Next consider reflection from a free end. To maintain the tension in the string at the free end, we connect the end to a light ring that can slide freely along a vertical rod (**Figure 16.21**). When it reaches a height equal

Figure 16.21 Reflection of a wave pulse at a boundary where the string end is free to move vertically.



CONCEPTS

to the maximum height of the pulse, however, the free end keeps moving upward, as you can see in Figure 16.21. The reason for this overshooting is that there are no string particles to the right of the free end exerting a downward force. As it overshoots, the free end exerts an upward force on the string to its left, pulling the string above the maximum height of the pulse. After reaching its maximum height, the ring moves back down to the equilibrium level. This up and down motion of the ring creates a left-moving pulse in the string, just as the up and down motion of the hand at the left created the original pulse. This leftward-moving pulse is the reflection of the original pulse. Although the reflected pulse is left–right reversed, it is not inverted.

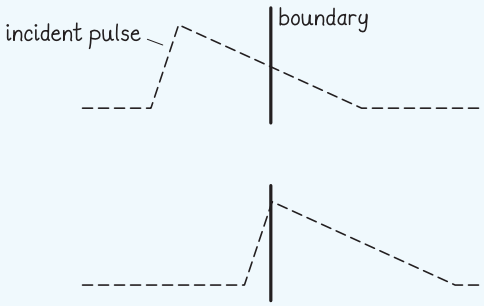
To construct the reflected pulse, we follow the same procedure we used to construct a pulse reflected from a fixed end, but we do not invert the reflected pulse.

EXAMPLE 16.2 Reflection from a free end

Let the triangular wave pulse of Figure 16.18 approach the free end of a string. Sketch the shape of the string (*a*) when a point halfway up the leading edge of the incident pulse has reached the fixed end and (*b*) when the peak of the pulse has reached the fixed end.

❶ **GETTING STARTED** I begin by sketching the pulse at the two instants given in the problem, ignoring the free end (**Figure 16.22**). I use a dashed line for the pulse to indicate that this is not the actual shape of the string at these instants, and I draw a vertical line to indicate the (horizontal) position of the free end of the string.

Figure 16.22

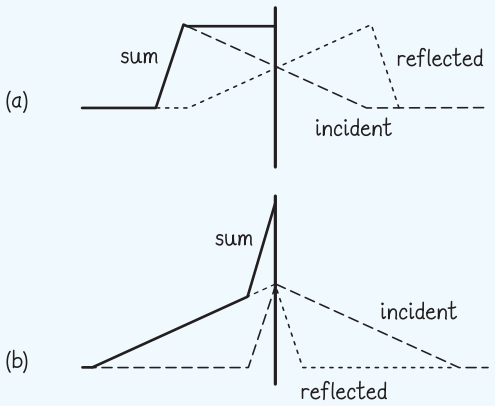


❷ **DEVISE PLAN** Once any part of the pulse reaches the free end, reflection begins. To determine the pulse shape during reflection, I must draw a reflected pulse that is reversed but not inverted and then add that reflected pulse to my incident pulses in Figure 16.22.

❸ **EXECUTE PLAN** (*a*) At the free end, the displacement of the reflected pulse is the same as that of the incident pulse, and so the string displacement is double what it was before the incident pulse arrived. In the region where the two pulses do not overlap, the sum of the displacements of the two pulses is equal to the displacement of the trailing edge of the incident pulse. Between this no-overlap region and the free end, I add the dashed and dotted lines algebraically to determine the shape of the string shown in **Figure 16.23*a***. ✔

(*b*) Following the same procedure yields Figure 16.23*b*. ✔

Figure 16.23



❹ **EVALUATE RESULT** Once the incident pulse arrives at the free end, the displacement of the free end at any given instant is twice that of the incident pulse, and therefore the free end swings far up in Figure 16.23*b*, in agreement with what I expect from Figure 16.21.

16.15 (*a*) When the free end of the string in Figure 16.21 reaches its maximum displacement at t_5, what is the kinetic energy in the pulse? (*b*) What is the potential energy in the pulse? (*c*) Is the energy in the pulse at this instant the same as the energy at t_1?

A case intermediate between a boundary where the end of the string is fixed and a boundary where it is free to move vertically is when a pulse reaches the boundary between two different mediums. The pulse is partially transmitted to the second medium and partially reflected. **Figure 16.24*a*** on the next page shows a heavy string connected to a lighter string. String 1 has a greater mass per unit length—a quantity called the **linear mass density** and represented by the symbol μ—than string 2. The wave speed along string 1 is c_1, and that along string 2 is c_2. The tensions in the two strings are equal, but because of the difference in mass density, $c_2 > c_1$. Because string 2 is much less massive than string 1, the end of string 1 behaves somewhat like a free end, and so the reflected pulse is not inverted.

CONCEPTS

Figure 16.24 Reflection and transmission of a wave pulse at the boundary between two strings that have different linear mass densities.

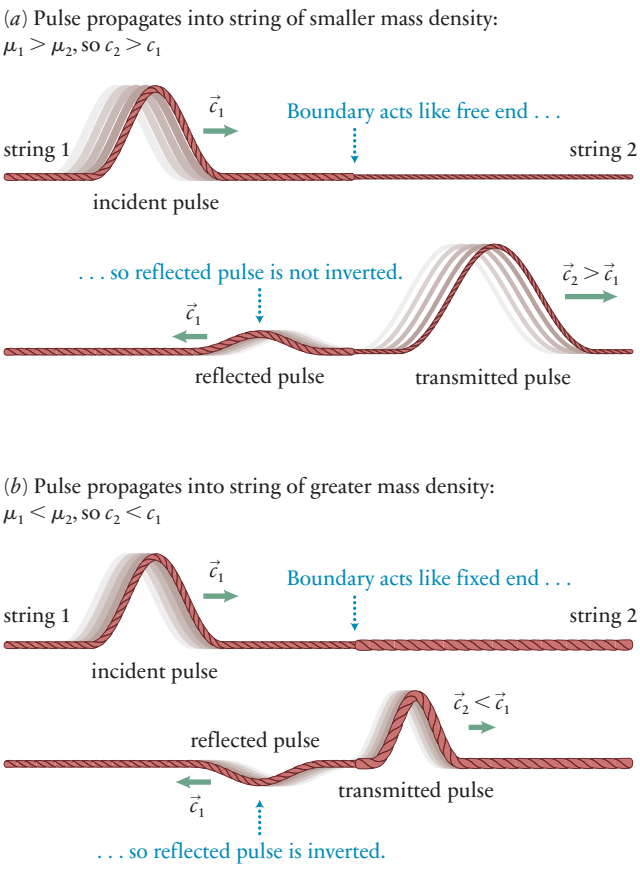


Figure 16.24*b* illustrates a pulse propagating along the same string 1 toward a string 2 that has a greater linear mass density. String 2 restricts the movement of the end of string 1, so now the boundary behaves like a fixed end, reflecting an inverted pulse back along string 1. The tensions in the two strings are again equal, but because of the difference in linear mass density, $c_2 < c_1$ and so the transmitted pulse travels more slowly than the incident pulse.

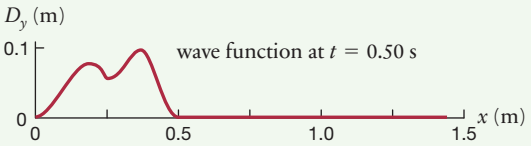
16.16 For the strings and pulse in Figure 16.24, what happens in the limit where (*a*) $\mu_2 \to 0$, (*b*) $\mu_2 \to \infty$, and (*c*) $\mu_2 = \mu_1$? (*d*) Why is the transmitted pulse in Figure 16.24*a* wider than the incident pulse?

CONCEPTS

Self-quiz

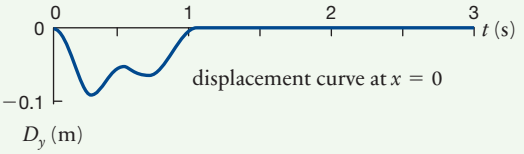
1. Given the wave function for $t = 0.50$ s shown in **Figure 16.25**, draw the displacement curves of particles at $x = 0$ and $x = 1.0$ m as the wave passes them. If the leading edge of the wave was at $x = 0$ at time $t = 0$, what is the speed of the wave?

Figure 16.25



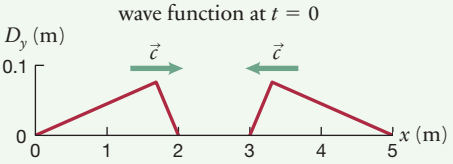
2. Given the displacement curve shown in **Figure 16.26** for a bead at $x = 0$ on a string as a wave moving toward the right with a speed of 2.0 m/s passes, draw the wave function at $t = 2.0$ s.

Figure 16.26



3. The two pulses shown in **Figure 16.27** are traveling in opposite directions; both have a speed of 1.0 m/s. Draw the sum of the two pulses at $t = 1.0$ s, $t = 2.0$ s, and $t = 3.0$ s.

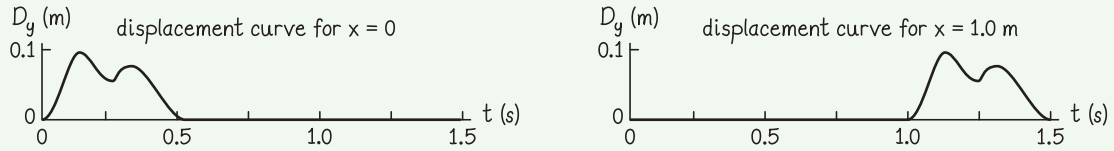
Figure 16.27



Answers

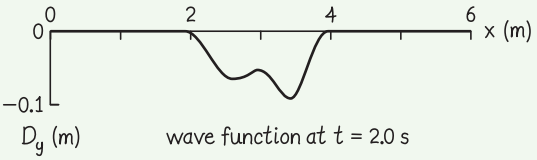
1. See **Figure 16.28**. The leading edge of the wave travels 0.50 m in 0.50 s, giving a wave speed of 1.0 m/s. In the displacement for the particle at $x = 0$, the leading edge passes $x = 0$ at $t = 0$ and the trailing edge passes $x = 0$ at 0.50 s. In the displacement curve for the particle at $x = 1.0$ m, the leading edge passes $x = 1.0$ m at $t = 1.0$ s and the trailing edge passes $x = 1.0$ m at 1.5 s.

Figure 16.28



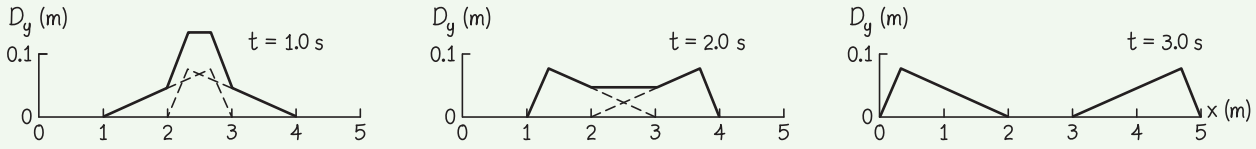
2. See **Figure 16.29**. With a wave speed of 2.0 m/s and 1.0 s needed for the whole wave to pass $x = 0$, the length of the wave must be 2.0 m. At $t = 1.0$ s, the leading edge of the wave passes $x = 2.0$ m. At $t = 2.0$ s, the leading edge passes $x = 4.0$ m and the trailing edge passes 2.0 m.

Figure 16.29



3. See **Figure 16.30**. The two pulses start out 1.0 m apart, and each has a width of 2.0 m. With a relative speed of 2.0 m/s, they begin to interfere at $t = 0.5$ s. At $t = 1.0$ s, the pulses overlap by half their width, with the trailing edges sticking out. At $t = 2.0$ s, the pulses also overlap by half their width, but now the leading edges are sticking out. At $t = 3.0$ s, the waves are separated again.

Figure 16.30



16.5 Wave functions

The displacement of the medium caused by a traveling wave is a function of both space and time, and so the mathematical description of a one-dimensional traveling wave requires a function in two variables. Let us begin by examining a traveling wave of constant shape, looking at the pulses first from a reference frame moving along with the wave and then from the Earth reference frame, which is stationary relative to the medium. In the moving reference frame, the wave is stationary and so its description is simple.

Figure 16.31*a* shows a transverse wave on a string at instant $t = 0$ when the origins of the two reference frames coincide. The wave travels in the positive x direction at speed c without changing its shape. Figure 16.31*b* shows the situation at some instant t, when the wave and the moving reference frame have traveled a distance ct to the right. As seen from the reference frame moving along with the wave, the wave remains stationary (Figures 16.31*c* and *d*). If the curve in the moving reference frame is described by some function $f(x_M)$, then the y component of the displacement $\vec{D}_M$ of any particle of the string as seen from the moving reference frame is given by

$$D_{My} = f(x_M). \tag{16.1}$$

What this tells us is that when we view the situation from the moving reference frame, x_M is the only variable we need to specify to determine the displacement D_{My}. The function $f(x_M)$, which specifies the (constant) shape of the wave as seen from the moving reference frame, is called the *time-independent wave function*.

When we view from the Earth reference frame, the wave has the same shape as in the moving reference frame but it is moving in the positive direction along the x axis. We must therefore specify both x and t in order to determine the displacement $\vec{D}$, and so we write $D_y = f(x, t)$. The function $f(x, t)$, called the *time-dependent wave function*, completely specifies

Figure 16.31 Traveling wave pulses seen from the Earth reference frame and from a reference frame moving along with the pulse at instant $t = 0$ (when the origins of the two reference frames overlap) and at some instant t later.

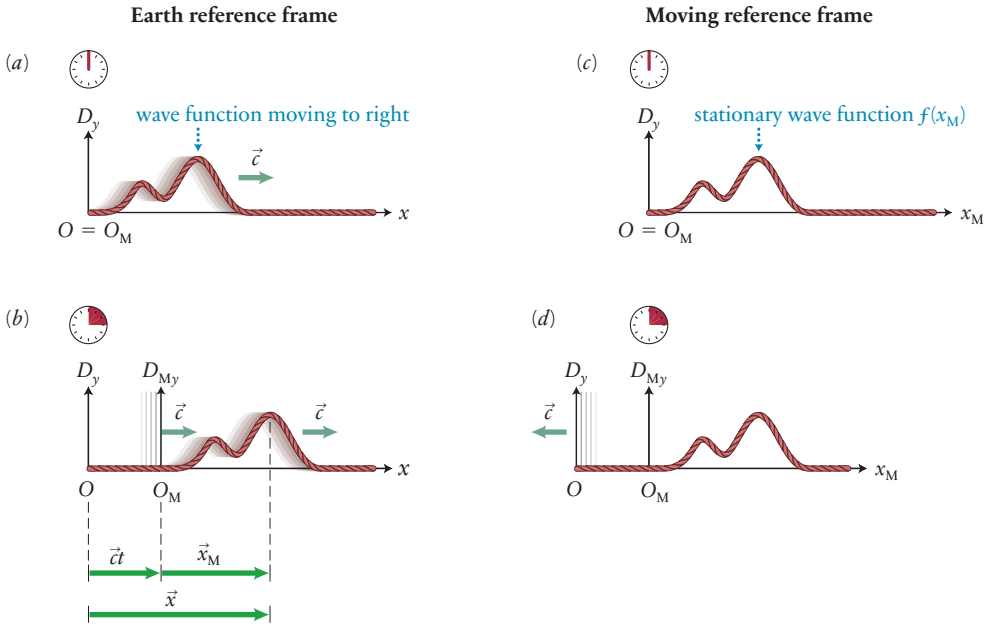
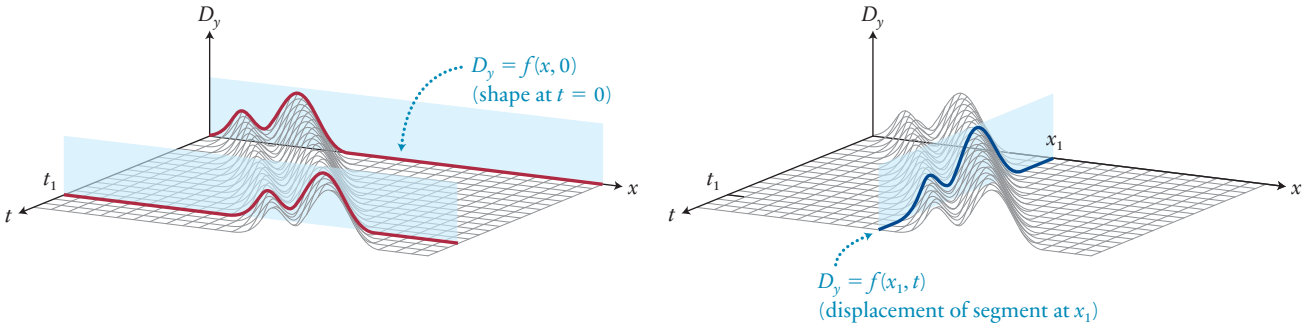


Figure 16.32 Time-dependent wave function.

(*a*) Shape of the medium at instants $t = 0$ and $t = t_1$

(*b*) Displacement of the medium at a fixed position as a function of time



the changing shape of the wave disturbance in the medium as seen from the Earth reference frame. To visualize this function, we need to make a three-dimensional graph of $f(x, t)$ versus the variables x and t (**Figure 16.32*a***). The surface shown in the graph gives the value of $D_y = f(x, t)$ for each pair of values of x and t. A vertical cut through the surface along the x axis yields the curve labeled $D_y = f(x, 0)$. This curve gives the shape of the wave disturbance at $t = 0$. A vertical cut through the surface parallel to the x axis intercepting the t axis at t_1 gives $D_y = f(x, t_1)$, the shape of the wave disturbance at t_1. As you can see, the wave has advanced in the positive x direction. Vertical cuts parallel to the t axis give the displacement at a fixed position in the medium as a function of time (Figure 16.32*b*). The curve $f(x_1, t)$, for example, shows how the displacement at x_1 changes with time.

The surface in Figure 16.32 is completely determined by the constant shape shown in Figure 16.31*c* and the wave speed c. To see how this is true, imagine sliding the curve in Figure 16.31*c* diagonally across the xt plane in Figure 16.32*a*. Because three-dimensional graphs are difficult to draw and to read, we shall use two-dimensional graphs to represent traveling waves that are constant in shape.

Let us now relate the displacement viewed from a moving reference frame to the displacement viewed from the Earth reference frame. At instant t, the y component of the displacement of any point P on the three-dimensional surface of an $f(x, t)$ graph is given by $D_y = f(x, t)$ in the Earth reference frame and by $D_{My} = f(x_M)$ in the moving reference frame (Figure 16.31*b*). Because the displacement of the string particles does not depend on which reference frame we use, we must have $D_y = D_{My}$. From Figure 16.31*b* we see that $x_M = x - ct$, so

$$f(x_M) = f(x - ct). \tag{16.2}$$

Substituting these relationships into Eq. 16.1, we discover that the y component of the displacement viewed from the Earth reference frame is given by

$$D_y = f(x - ct). \tag{16.3}$$

This equation describes a wave of constant form $f(x)$ traveling in the positive x direction at speed c. For a wave traveling in the negative x direction at speed c, we have

$$D_y = f(x + ct). \tag{16.4}$$

Any wave traveling along the x axis has either the form $f(x - ct)$ or the form $f(x + ct)$. Provided the quantities x, c, and t appear in the combination $(x - ct)$ or $(x + ct)$ in a function, that function can represent a traveling wave. Thus, the function $\sin[k(x - ct)]$, where k is a constant, represents a harmonic wave that travels in the positive direction along the x axis. The function $\sin[k^2(x^2 - c^2t^2)]$, however, does not represent a traveling harmonic wave.

EXERCISE 16.3 Traveling waves

(a) Consider the time-dependent wave function

$$D_y = f(x, t) = \begin{cases} b(x - ct) & \text{for } 0 < x - ct \leq 1.0 \text{ m} \\ 0 \text{ for} & x - ct \leq 0 \quad \text{or} \quad x - ct > 1.0 \text{ m}, \end{cases}$$

where $b = 0.80$ and $c = 2.0$ m/s. Plot the time-independent wave function for a few values of t to verify that the function corresponds to a wave traveling in the positive x direction at a speed of 2.0 m/s.

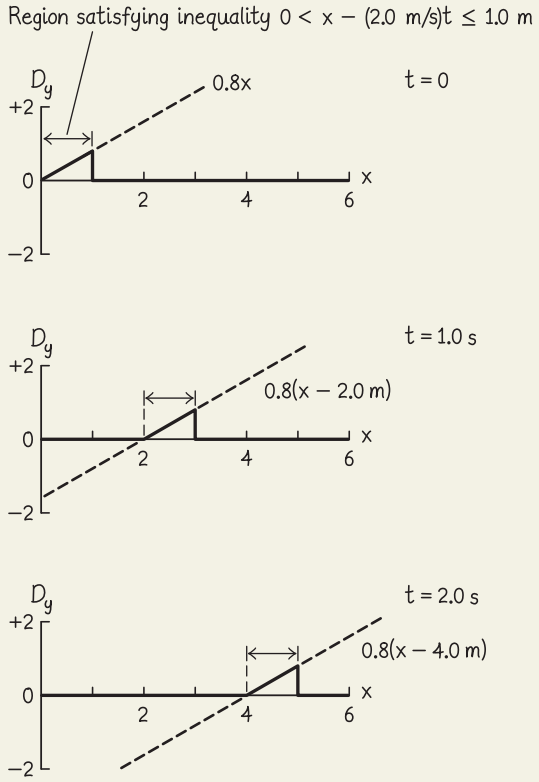
(b) Let the shape of a wave at $t = 0$ be described by the function

$$f(x, 0) = \frac{a}{x^2 + b}.$$

If the wave travels in the negative x direction at a speed c, what is the mathematical form of the time-dependent wave function $f(x, t)$?

SOLUTION (a) Because I need to plot the function versus x, I first add ct to each term in the inequality $0 < x - ct \leq 1.0$ m and get $ct < x \leq (1.0 \text{ m}) + ct$. For $t = 0$, the function is nonzero when $0 < x \leq 1.0$ m; for $t = 1.0$ s when $2.0 \text{ m} < x \leq 3.0$ m; for $t = 2.0$ s when $4.0 \text{ m} < x \leq 5.0$ m; and so on. So, for $t = 0$ I plot the function $(0.80)x$ between $x = 0$ and $x = 1.0$ m; for $t = 1.0$ s I plot the function $0.80(x - 2.0 \text{ m})$ between $x = 2.0$ m and $x = 3.0$ m; and so on (**Figure 16.33**). My graphs show a triangular wave that is constant in shape and displaced in the positive x direction by +2.0 m each second. ✔

Figure 16.33



QUANTITATIVE TOOLS

(*b*) When viewed from a reference frame traveling along with the wave, the wave function does not depend on time and has the form

$$f(x_M) = \frac{a}{x_M^2 + b}.$$

Equation 16.2 gives me the relationship between the displacement viewed from the Earth reference frame and one moving along with the pulse in the positive x direction. Because the pulse in part *b* of this exercise moves in the negative direction, I have to invert the sign in front of the speed c, so $f(x + ct) = f(x_M)$ and $x_M = x + ct$. Substituting these relationships into $f(x_M)$, I obtain the time-dependent wave function

$$f(x, t) = \frac{a}{(x + ct)^2 + b}. \checkmark$$

Let us consider a transverse harmonic wave traveling along a string aligned with an x axis. Such a wave can be generated by making the end of the string execute a simple harmonic motion. **Figure 16.34*a*** shows the shape of the string at an instant at which the left end, at $x = 0$, is in the equilibrium position. For convenience, we choose this instant to be $t = 0$. The time-independent wave function is

$$f(x) = A \sin(kx), \tag{16.5}$$

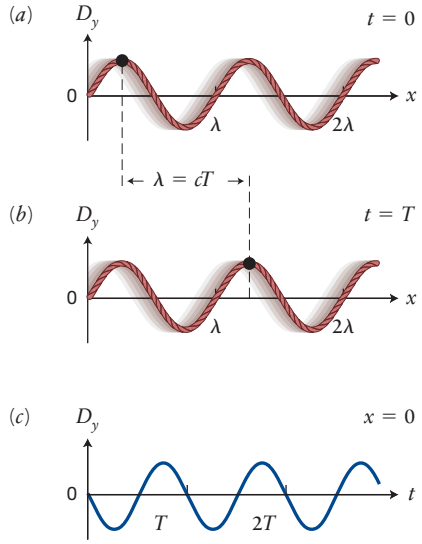
where A is the amplitude of the wave and k is a constant.The wave repeats itself over a distance λ (the wavelength), so the displacement $f(x)$ is the same at both ends of a wavelength—that is, at x and $x + \lambda$:

$$A \sin(kx) = A \sin(kx + k\lambda). \tag{16.6}$$

Because a sine function repeats itself when its argument is increased by 2π, it follows from Eq. 16.6 that $k\lambda = 2\pi$, so

$$k = \frac{2\pi}{\lambda}. \tag{16.7}$$

Figure 16.34 Wave function of a harmonic wave (*a*) at $t = 0$ and (*b*) after one period. (*c*) The corresponding displacement curve for $x = 0$.



The constant k, called the **wave number** of the motion, thus gives the number of wavelengths in a length of 2π m. Because wavelength has units of meters, the wave number k has SI units of m^{-1}.

If the wave travels at speed c in the positive x direction, the time-dependent wave function is obtained by replacing x in Eq. 16.5 by $(x - ct)$:

$$D_y = f(x, t) = A \sin[k(x - ct)]. \tag{16.8}$$

Figure 16.34*b* shows the wave function after the string end at $x = 0$ has executed one complete cycle (that is, at $t = T$, where T is the period of the motion). During this cycle, the wave has advanced by a distance corresponding to one wavelength. To see what this means, compare the positions of the black dots in Figures 16.34*a* and *b*. Think of these two dots as representing the same point on the traveling wave. In the time interval from $t = 0$ to $t = T$, this point of the wave has moved a distance λ. (Do *not* think of the dots as representing the same particle *of the string* because, remember, the medium does not move horizontally in this wave motion; only the momentum and energy associated with the wave move horizontally.) Because the wave moves at wave speed c, this distance corresponds to

$$\lambda = cT \tag{16.9}$$

or, substituting the frequency $f \equiv 1/T$,

$$\lambda f = c. \tag{16.10}$$

Because the period of the motion is T, we have for $x = 0$ $\sin(kct) = \sin[kc(t + T)]$. This can be true only if $kcT = 2\pi$, or

$$kc = \frac{2\pi}{T} \equiv \omega, \tag{16.11}$$

where ω is the angular frequency as defined in Section 15.5: $\omega = 2\pi f = 2\pi/T$. Using the angular frequency, we can rewrite the time-dependent wave function in Eq. 16.8 in the form:

$$D_y = f(x, t) = A \sin(kx - \omega t). \tag{16.12}$$

Let us now analyze the vertical displacement of the string as the wave passes through. We can use Eq. 16.12 to examine the vertical displacement of any particle of the string as a function of time. For instance, substituting $x = 0$ into Eq. 16.12 and using the trigonometric identity $\sin(-\alpha) = -\sin\alpha$, we obtain

$$D_y = f(0, t) = A \sin(-\omega t) = -A \sin(\omega t). \tag{16.13}$$

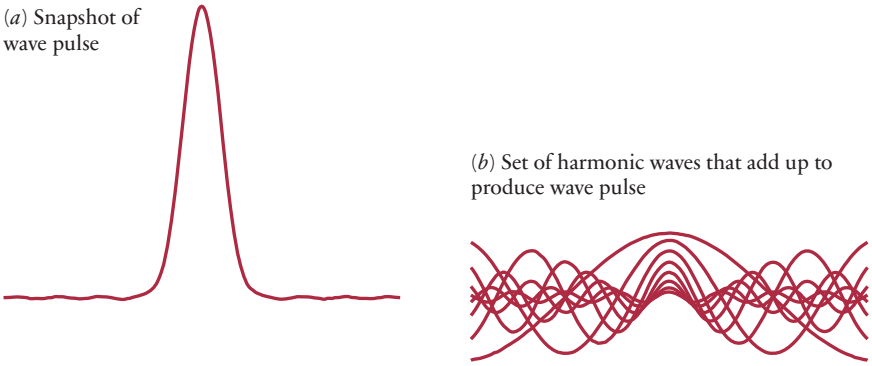
The displacement curve for this function $D_y(0, t)$ is shown in Figure 16.34*c*. (If you are having trouble understanding why the displacement curve dips downward initially, think about what is going to happen in Figure 16.34*a* to the string particle located at $x = 0$ just after $t = 0$: It is going to be pulled down below the x axis.)

Next we consider a point a distance d from the end of the string. Substituting $x = d$ into Eq. 16.12 yields

$$\begin{aligned} D_y = f(d, t) &= A \sin(kd - \omega t) \\ &= A \sin(-\omega t + kd) \end{aligned} \tag{16.14}$$

QUANTITATIVE TOOLS

Figure 16.35 Any wave can be expressed in terms of sinusoidally varying (harmonic) waves.



or, using the trigonometric identity $\sin(-\alpha) = \sin(\alpha - \pi)$,

$$D_y = f(d, t) = A \sin(\omega t - kd - \pi) = A \sin(\omega t + \text{constant}). \quad (16.15)$$

This expression, which gives the y component of the displacement of the string at any arbitrary position $x = d$, shows that all points along the string execute a simple harmonic motion with a period $2\pi/\omega = T$.

In deriving Eq. 16.8, we assumed that $f = 0$ at $x = 0$ and $t = 0$. To obtain a more general expression that allows for a nonzero initial displacement, we add to the argument of the sine in Eq. 16.12 (the *phase* of the wave, see Section 15.5) an initial phase ϕ_i:

$$D_y = f(x, t) = A \sin(kx - \omega t + \phi_i). \quad (16.16)$$

The initial phase specifies the initial condition: At $x = 0$ and $t = 0$, Eq. 16.16 becomes $D_y = f(0, 0) = A \sin \phi_i$, and so, for a given amplitude, ϕ_i determines the initial value of the vertical displacement at $x = 0$.

In Section 15.3 we discussed Fourier's theorem, which states that the time dependence of any function can be expressed in terms of sinusoidally varying functions of time. Because of the superposition principle, the same theorem can be applied to waves: Any wave can be expressed in terms of sinusoidally varying waves (harmonic waves). **Figure 16.35a**, for example, shows a wave pulse obtained by adding together the harmonic waves shown in Figure 16.35*b*. At the center, all of the harmonic waves add up to form the pulse; everywhere else, they add up to nearly zero (surprising, but true!). So, if we understand harmonic waves, we can deal with more complicated waves by determining the harmonic waves that add up to give these waves.



16.17 (*a*) Which of the following functions could represent a traveling wave?

(*i*) $A \cos(kx + \omega t)$

(*ii*) $e^{-k|x-ct|^2}$

(*iii*) $b(x - ct)^2 e^{-x}$

(*iv*) $-(b^2 t - x)^2$

(*b*) Which of the following functions can be made to represent a traveling wave?

(*i*) $x/(1 + bx^2)$

(*ii*) xe^{-kx}

(*iii*) x^2

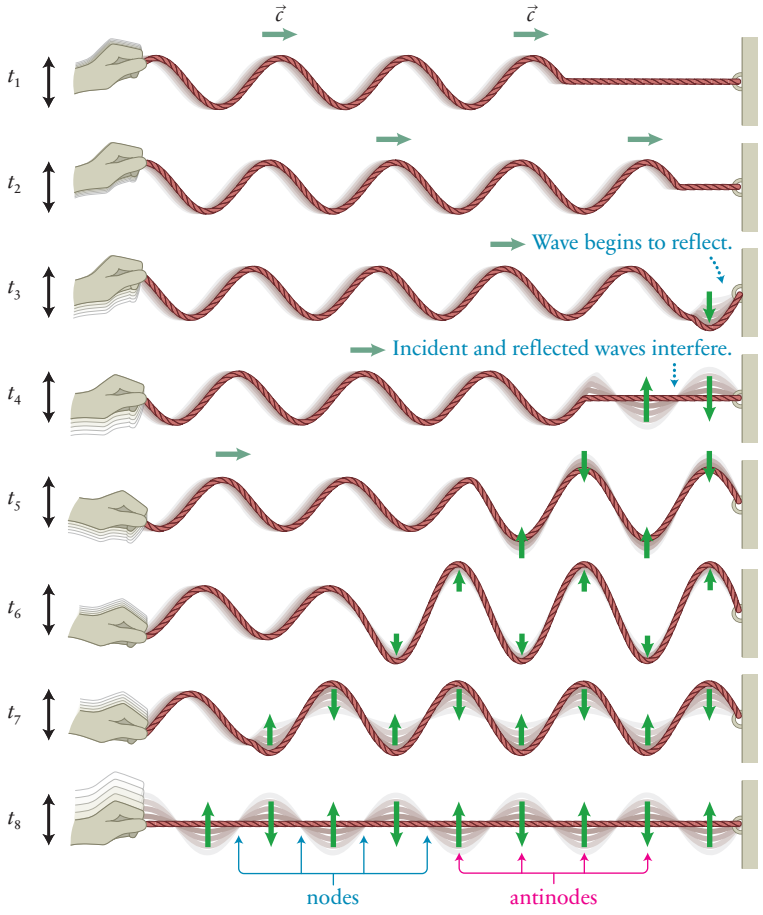


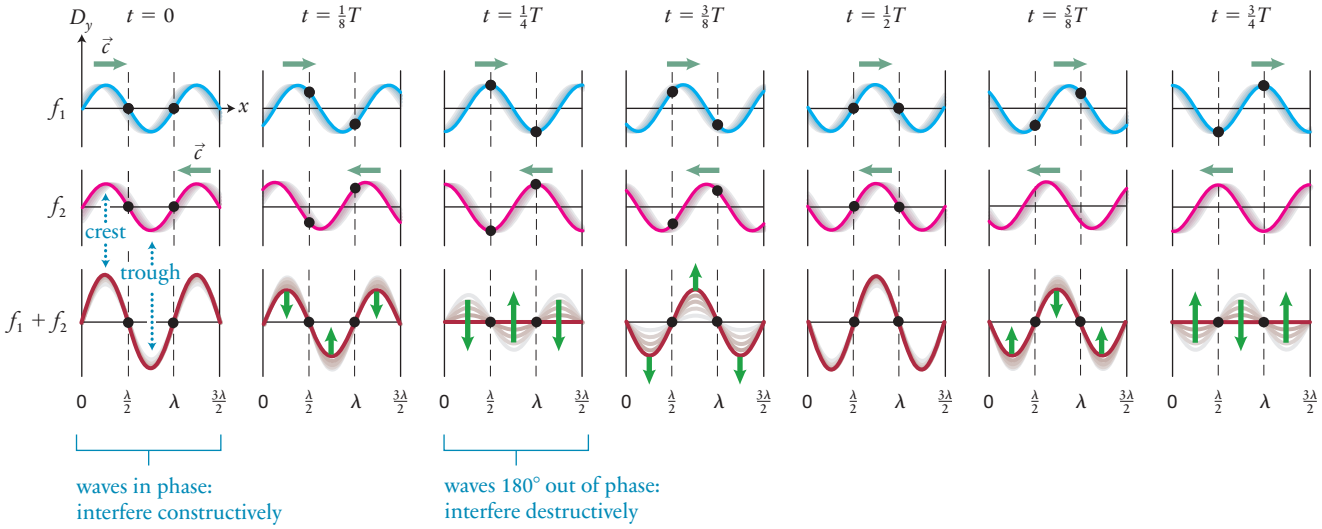
Figure 16.36 When a harmonic wave is reflected from the fixed end of a string, the reflected wave interferes with the incident wave, forming a standing wave.

16.6 Standing waves

When a harmonic wave travels along a string that has a fixed end, the reflected wave interferes with the incident wave. **Figure 16.36** shows the striking pattern that results from this interference. Initially the wave travels to the right, but as the reflected wave begins traveling back at t_3, the two waves traveling in opposite directions create a pulsating *stationary* pattern. Some particles of the string—the nodes (see Section 16.3)—do not move at all, while the sinusoidal loops between the nodes move up and down rather than along the string. The motion has its greatest amplitude at points halfway between the nodes; these points are called **antinodes**. The pulsating stationary pattern caused by harmonic waves of the same amplitude traveling in opposite directions is called a **standing wave.**

Figure 16.37 illustrates how standing waves come about. Each diagram shows a blue wave propagating to the right and a red wave propagating to the left. The two waves have the same amplitude and wavelength. The speed at which they propagate is determined by the medium and is therefore the same for both waves. The bottom part of each diagram shows the sum of the two waves, obtained by adding, for each value of x, the vertical displacements of the two waves at that value of x. In the diagram labeled $t = 0$, the crests of the red wave have the same x coordinate as the crests of the blue wave, and the same is true for the troughs: Each trough in the red wave aligns with a trough in the blue wave. Waves like this, where comparable points align, are said to be *in phase* (their phases are the same). The two waves reinforce each other—they interfere constructively—and adding them yields a wave that has twice the amplitude. A quarter period later, at $t = \frac{1}{4}T$, one has a crest when the other has a trough and the waves

Figure 16.37 Standing wave created by two counterpropagating waves f_1 and f_2.



are 180° *out of phase*. The two waves now interfere destructively, and the string is flat at this instant.

16.18 (*a*) In the standing wave pattern of Figure 16.37, how is the energy distributed between kinetic and potential at $t = 0$, $t = \frac{1}{8}T$, and $t = \frac{1}{4}T$? (*b*) Is the energy in a length of the string corresponding to one wavelength constant? (*c*) Does the standing wave transport energy? If so, in which direction? If not, why not?

To see why certain points are nodes, let us denote the y components of the displacements of the red and blue waves by $D_{1y} = f_1(x, t)$ and $D_{2y} = f_2(x, t)$, respectively, and look at what happens at the points $x = \lambda/2$ and $x = \lambda$. At $t = 0$, we have $D_{1y} = D_{2y} = 0$, and so $D_{1y} + D_{2y} = 0$. Now consider the situation at $t = \frac{1}{8}T$. Wave f_1 has advanced to the right, while wave f_2 has moved an equal distance to the left. At $\lambda/2$, the vertical displacement of the string caused by wave f_1 is now positive, but that caused by wave f_2 is negative and equal in magnitude, so the two add to zero and the string does not move from its equilibrium position. At $x = \lambda$ in the $t = \frac{1}{8}T$ diagram, we see the same thing but with the displacement directions reversed: f_1 displaces the string in the negative y direction, and f_2 displaces it by the same amount in the positive y direction. Because both waves always travel equal distances in opposite directions, the displacements caused by them at the points $x = \lambda/2$ and $x = \lambda$ are *always* equal in magnitude and opposite in sign. Consequently $D_{1y} + D_{2y}$ remains zero at these points for all t.

We can analyze the situation quantitatively by writing the wave traveling to the right as

$$D_{1y} = f_1(x, t) = A \sin(kx - \omega t) \tag{16.17}$$

and the wave traveling to the left as

$$D_{2y} = f_2(x, t) = A \sin(kx + \omega t). \tag{16.18}$$

The combined wave then is

$$D_y = f_1(x, t) + f_2(x, t) = A[\sin(kx - \omega t) + \sin(kx + \omega t)]. \tag{16.19}$$

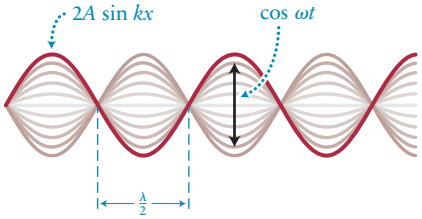


Figure 16.38 A standing wave is a stationary harmonic wave with an amplitude that varies like a sine. The spacing between adjacent nodes is equal to $\lambda/2$.

Using the trigonometric identities $\sin\alpha+\sin\beta = 2\sin\frac{1}{2}(\alpha+\beta)\cos\frac{1}{2}(\alpha-\beta)$ and $\cos\alpha = \cos(-\alpha)$ (see Appendix B), we can write Eq. 16.19 as

$$D_y = f_1(x,t) + f_2(x,t) = 2A\sin kx\cos\omega t = [2A\sin kx]\cos\omega t. \quad (16.20)$$

This is *not* a traveling wave because it is not a function of $x-ct$ or of $x+ct$ (see the text immediately following Eq. 16.4). Instead this function describes a standing wave on a string, one in which the string displacement oscillates harmonically as $\cos\omega t$ with an amplitude that varies along x as $2A\sin kx$ (**Figure 16.38**).

For a standing wave of the form given in Eq. 16.20 the displacement is *always* zero at the locations where $\sin kx$ is zero. These points are the nodes; they occur when kx is a whole-number multiple of π. Using the fact that $k = 2\pi/\lambda$ (Eq. 16.7), we obtain

$$\frac{2\pi}{\lambda}x = n\pi \qquad n = 0, \pm1, \pm2, \ldots, \quad (16.21)$$

and so nodes occur at

$$x = 0, \pm\frac{\lambda}{2}, \pm\lambda, \pm\frac{3\lambda}{2}, \ldots. \quad (16.22)$$

The displacement has its maximum value at the points where $\sin kx = \pm1$. These antinodes occur when kx is an odd whole-number multiple of $\pi/2$:

$$\frac{2\pi}{\lambda}x = \frac{n\pi}{2} \qquad n = \pm1, \pm3, \pm5, \ldots, \quad (16.23)$$

and so antinodes occur at

$$x = \pm\frac{\lambda}{4}, \pm\frac{3\lambda}{4}, \pm\frac{5\lambda}{4}, \ldots. \quad (16.24)$$

Note that, as shown in Figure 16.38, adjacent nodes are separated by a distance $\lambda/2$. Adjacent antinodes are also spaced by a distance $\lambda/2$ and lie midway between the nodes.

16.19 (*a*) Do two counterpropagating waves that have the same wavelength but different amplitudes cause standing waves? (*b*) Do two counterpropagating waves that have the same amplitude but different wavelengths cause standing waves?

16.7 Wave speed

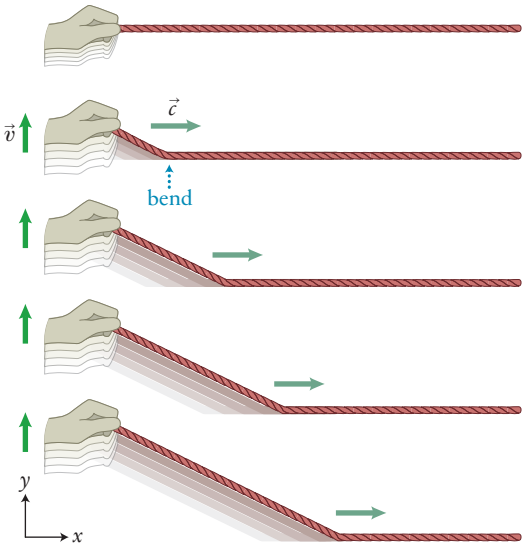
To determine a quantitative expression for the wave speed c, consider the situation shown in **Figure 16.39**. The end of a taut horizontal string is moved vertically upward at a constant velocity $\vec{v}$. The displacement of the end causes a triangular wave pulse to propagate along the string. The bend on the leading edge travels in the horizontal direction with wave speed c. Just as with the string of beads in Section 16.2, all segments of the string execute the same motion as the end of the string: Each segment begins at rest and then, as the pulse reaches it, moves vertically upward at speed v.

Let us consider the motion of the string during a small time interval Δt beginning at some instant t. At the beginning of the interval, the end of the string has been raised a distance vt and the bend has traveled a distance ct (Figure 16.39*b*). The slanted segment of the string moves vertically up at constant speed v; at the instant represented in Figure 16.39*b*, segment A moves up while the rest of the string remains horizontal. A short time

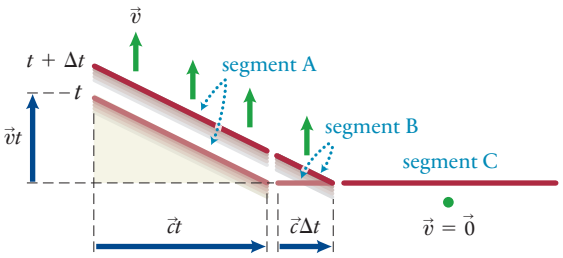
QUANTITATIVE TOOLS

Figure 16.39 Procedure for obtaining a quantitative expression for the wave speed c.

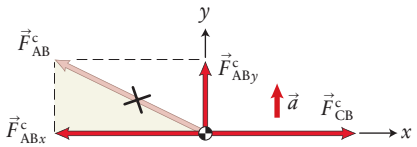
(a) Triangular wave pulse generated by lifting end of string at constant velocity



(b) In time interval Δt, bend advances distance $c\Delta t$



(c) Free-body diagram for segment B



interval Δt later, segment A has moved farther up and is still moving at v, so the momentum of segment A has not changed.

Let us choose the x axis in the direction of the wave and the y axis in the upward direction. During the time interval Δt, the bend has advanced by a distance $\Delta x = c\Delta t$, and so an additional length $c\Delta t$ of string (segment B) now also moves vertically upward with speed v. The mass of this segment is $\mu(c\Delta t)$, where μ is the mass per unit length, or **linear mass density,** of the string. For a uniform string of length ℓ and mass m, the linear mass density is

$$\mu \equiv \frac{m}{\ell} \quad \text{(uniform linear object)}. \tag{16.25}$$

The SI units of linear mass density are kg/m. The change in the momentum of segment B is equal to the mass of B times its change in velocity $\Delta\vec{v}_B = \vec{v} - 0 = \vec{v}$. The mass of B is equal to the linear mass density of the string times the equilibrium length of B, $m_B = \mu(c\Delta t)$, and so $\Delta\vec{p}_B = m_B\Delta\vec{v}_B = \mu(c\Delta t)\vec{v}$. This change in momentum occurs during the time interval Δt, and so the rate of change is

$$\frac{\Delta\vec{p}_B}{\Delta t} = \frac{\mu(c\Delta t)\vec{v}}{\Delta t} = \mu c\vec{v}. \tag{16.26}$$

The rate of change in momentum of segment B of the string is by definition equal to the vector sum of the forces exerted on it:

$$\Sigma\vec{F}_B = \frac{\Delta\vec{p}_B}{\Delta t} = \mu c\vec{v}. \tag{16.27}$$

Figure 16.39c shows the free-body diagram for segment B. Segment A exerts on B a force slanting upward along the direction of segment A; segment C exerts on B a force in the horizontal direction. The x components of these two forces must balance because segment B is not accelerating in the horizontal direction. We also know that the x components of these two forces are equal in magnitude to the tension $\mathcal{T}$ in the string. The

vector sum of the forces exerted on B is thus equal to the vertical component of the force exerted by A on B. Finally, we note that the shaded triangles in Figures 16.39*b* and *c* are similar, so

$$\frac{F^c_{ABy}}{F^c_{ABx}} = \frac{F^c_{ABy}}{\mathcal{T}} = \frac{vt}{ct} = \frac{v}{c}, \tag{16.28}$$

and thus the vector sum of the forces acting on segment B of the string is

$$\Sigma F_y = F^c_{ABy} = \frac{\mathcal{T}v}{c}. \tag{16.29}$$

Combining this result with Eq. 16.27, we get $\mu cv = \mathcal{T}v/c$, or

$$c = \sqrt{\frac{\mathcal{T}}{\mu}}. \tag{16.30}$$

Note that the wave speed is independent of the velocity v of the string segments. The result in Eq. 16.30 agrees with the qualitative arguments in Section 16.2: The speed of a pulse along a string increases with increasing tension $\mathcal{T}$ and with decreasing linear mass density.

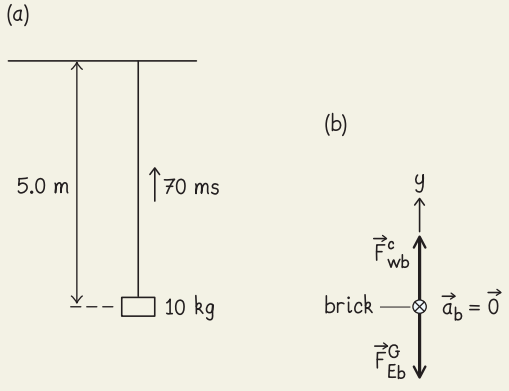
16.20 (*a*) Draw a free-body diagram for segment A of the string at instant *t* in Figure 16.39. What is the vector sum of the forces exerted on A? (*b*) What does your answer to part *a* tell you about the change in momentum of A? Does your conclusion about A's momentum make sense? (*c*) Does the hand do work on the string? (*d*) What happens to the energy of the string as the end moves upward?

EXAMPLE 16.4 Measuring mass

You use a hammer to give a sharp horizontal blow to a 10-kg lead brick suspended from the ceiling by a wire that is 5.0 m long. It takes 70 ms for the pulse generated by the sudden displacement of the brick to reach the ceiling. What is the mass of the wire?

1 GETTING STARTED I begin by making a sketch of the situation (**Figure 16.40*a***). Because I know the pulse travels 5.0 m in 70 ms, I can determine the wave speed along the wire: $c = (5.0\text{ m})/(0.070\text{ s}) = 71\text{ m/s}$. This wave speed depends on the linear mass density of the wire and the tension in the wire. The latter is determined by the force exerted by the brick on the wire. Because the force exerted by the brick on the wire and the force exerted by the wire on the brick are equal in magnitude, I draw a free-body diagram for the brick (Figure 16.40*b*) to help me determine that magnitude.

Figure 16.40



2 DEVISE PLAN To determine the mass of the wire, I can use Eq. 16.25. To use this equation, I must know the length of the wire, which is given, and the linear mass density μ, which I can calculate from Eq. 16.30. To obtain μ from Eq. 16.30, I need to know c (which I already calculated) and the tension $\mathcal{T}$ in the wire. This tension is equal to F^c_{bw}, which in turn is equal to F^c_{wb}. From my free-body diagram, I see that $F^c_{wb} = F^G_{Eb}$, and so I know F^c_{bw} and therefore $\mathcal{T}$.

3 EXECUTE PLAN The tension is equal to the downward force of gravity exerted on the brick: $\mathcal{T} = F^G_{Eb} = (10\text{ kg})(9.8\text{ m/s}^2) = 98\text{ N}$. Substituting the values for $\mathcal{T}$ and c into Eq. 16.30, I calculate the linear mass density of the wire $\mu = \mathcal{T}/c^2 = (98\text{ N})/(71\text{ m/s})^2 = 0.019\text{ kg/m}$. The mass of the wire is thus $m = (0.019\text{ kg/m})(5.0\text{ m}) = 9.5 \times 10^{-2}\text{ kg}$. ✔

4 EVALUATE RESULT The value I obtain—about 0.1 kg—is not unreasonable for a 5.0-m-long wire that can support a lead brick.

QUANTITATIVE TOOLS

16.8 Energy transport in waves

Imagine moving the end of a very long, taut string up and down so that a harmonic wave begins traveling along the string at the wave speed c given in Eq. 16.30. Energy moves along with the wave, and so, as you shake the string end, you must supply energy to the string. At what rate must you supply this energy? To answer this question, let us first consider the energy involved in generating a simple triangular wave as shown in **Figure 16.41**.

As it raises the end of the string, the hand does work on the string. The work done by the hand on the string during a time interval Δt is equal to the scalar product of the force $\vec{F}^{\text{c}}_{\text{hA}}$ exerted by the hand on segment A of the string and the force displacement. From Figures 16.41*a* and *b*, we see that

$$W = F^{\text{c}}_{\text{hA}y}(v\Delta t), \tag{16.31}$$

where $F^{\text{c}}_{\text{hA}y}$ is the y component of $\vec{F}^{\text{c}}_{\text{hA}}$ and $v\Delta t$ is the y component of the force displacement. If we ignore the force of gravity on segment A of the string, the force exerted by the hand on A is transmitted undiminished to B, and so $\vec{F}^{\text{c}}_{\text{hA}}$ is equal in magnitude to the force $\vec{F}^{\text{c}}_{\text{AB}}$ exerted by A on B. Using Eq. 16.29, we then have

$$F^{\text{c}}_{\text{hA}y} = \mathcal{T}\frac{v}{c}. \tag{16.32}$$

Substituting this result and $\mathcal{T} = \mu c^2$ from Eq. 16.30 into Eq. 16.31, we discover that the work done by the hand on the string during the time interval Δt is

$$W = \left(\mathcal{T}\frac{v}{c}\right)v\Delta t = \mu c\, v^2\Delta t. \tag{16.33}$$

The energy law, $\Delta E = W$ (Eq. 9.1), tells us that the work done by the hand on the string changes the energy of the string. The average rate at which the hand supplies energy to the string (the *average power*) is equal to the change in the energy of the string divided by the time interval over which the energy changes (Eq. 9.29), and so $P_{\text{av}} \equiv \Delta E/\Delta t = \mu c v^2$.

As the hand does work on the string, the kinetic energy and the potential energy of the string change: The kinetic energy changes because more of the string is set in motion, and the potential energy changes because the raising stretches the string, storing elastic potential energy. Let us first calculate the change in kinetic energy. After a time interval Δt, segment A has a velocity v, and so its kinetic energy is $\frac{1}{2}m_{\text{A}}v^2$, where m_{A} is the mass of A. The original length of segment A is $c\Delta t$ (Figure 16.41*a*), and so $m_{\text{A}} = \mu c\Delta t$. The change in kinetic energy is thus

$$\Delta K = \tfrac{1}{2}(\mu c\Delta t)v^2 - 0 = \tfrac{1}{2}\mu c v^2\Delta t, \tag{16.34}$$

which, Eq. 16.33 tells us, is exactly half the energy supplied to the string!

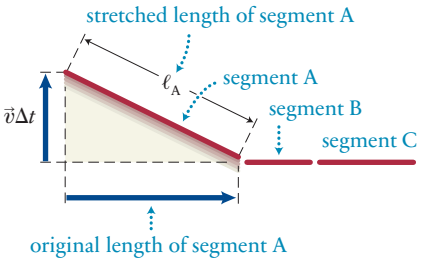
The other half must be potential energy associated with the stretching of segment A (the rest of the string is undisturbed). Let the length of A at the instant shown in Figure 16.41*a* be ℓ_{A}. The change in the length of the string is then $\Delta\ell = \ell_{\text{A}} - c\Delta t$. If $\Delta\ell$ is small, we can treat the string as a spring of spring constant k_{s}, and so Eq. 9.23 tells us the change in potential energy stored in the stretched string is

$$\Delta U = \tfrac{1}{2}k_{\text{s}}\Delta\ell^2 = \tfrac{1}{2}(k_{\text{s}}\Delta\ell)\Delta\ell = \tfrac{1}{2}F^{\text{c}}_{\text{hA}}\Delta\ell, \tag{16.35}$$

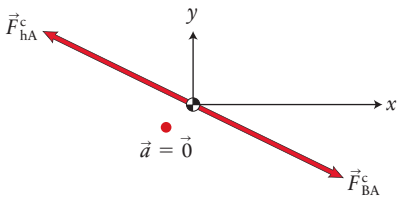
where we have used Hooke's law, Eq. 8.18, to substitute $F^{\text{c}}_{\text{hA}} = k_{\text{s}}\Delta\ell$. Because segment A is not accelerating, we see from the free-body diagram

Figure 16.41 Procedure for determining the energy involved in generating a simple triangular wave.

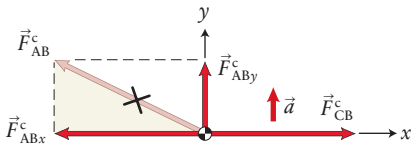
(*a*) Triangular wave pulse generated by lifting end of string at constant velocity



(*b*) Free-body diagram for segment A



(*c*) Free-body diagram for segment B



QUANTITATIVE TOOLS

for segment A (Figure 16.41*b*) that $\vec{F}^{\,c}_{hA} = -\vec{F}^{\,c}_{BA}$. Because the force exerted by segment B on segment A and the force exerted by segment A on segment B form an interaction pair, we have $\vec{F}^{\,c}_{hA} = \vec{F}^{\,c}_{AB}$. Comparing the similar shaded triangles in Figure 16.41*a* and the free-body diagram for segment B (Figure 16.41*c*), we see that

$$\frac{F^c_{AB}}{F^c_{ABx}} = \frac{F^c_{hA}}{\mathcal{T}} = \frac{\ell_A}{c\Delta t} \tag{16.36}$$

or $F^c_{hA} = \mathcal{T}\ell_A/(c\Delta t)$, and so Eq. 16.35 becomes

$$\Delta U_A = \tfrac{1}{2}\frac{\mathcal{T}\ell_A}{c\Delta t}\Delta\ell = \tfrac{1}{2}(\mu c^2)\frac{\ell_A}{c\Delta t}\Delta\ell = \tfrac{1}{2}\frac{\mu c}{\Delta t}\ell_A\Delta\ell, \tag{16.37}$$

where we have used Eq. 16.30 to eliminate $\mathcal{T}$. As long as the vertical displacement of the end of the string is small, $\ell_A \approx c\Delta t$, and so

$$\ell_A\Delta\ell = \ell_A(\ell_A - c\Delta t) = \ell_A^2 - \ell_A(c\Delta t) \approx \ell_A^2 - (c\Delta t)^2 \tag{16.38}$$

or, when we apply the Pythagorean theorem to the shaded triangle in Figure 16.41*a*, $\ell_A\Delta\ell = (v\Delta t)^2$, and therefore

$$\Delta U = \tfrac{1}{2}\frac{\mu c}{\Delta t}(v\Delta t)^2 = \tfrac{1}{2}\mu c v^2\Delta t. \tag{16.39}$$

So, in the limit of small displacement, the change in elastic potential energy is indeed equal to the change in kinetic energy, and the triangular pulse contains equal amounts of kinetic and potential energy.

Let us next consider the energy carried in a harmonic wave described by $y(x, t) = A\sin(kx - \omega t)$. We begin by calculating the amount of energy E_λ stored in a length of string corresponding to one wavelength of the harmonic wave. Because during each period T the wave advances by one wavelength, E_λ is the amount of energy that must be supplied in one period. The average power is then

$$P_{av} \equiv \frac{\Delta E}{\Delta t} = \frac{E_\lambda}{T}. \tag{16.40}$$

To calculate E_λ, we first divide the string into very small segments, each of length dx. If the linear mass density of the string is μ, the mass of each segment is $dm = \mu\, dx$. Each segment carries out a simple harmonic motion, and the energy associated with this simple harmonic motion is $\frac{1}{2}(dm)\omega^2A^2$ (Eq. 15.17). At any given instant, part of this energy is in the form of kinetic energy associated with the motion of the segment and part is in the form of potential energy associated with its deformation. The sum of the two, however, always equals $\frac{1}{2}(dm)\omega^2A^2$. All the segments in a length of string one wavelength long add up to a mass of $\sum dm = m$, and from the definition of linear mass density, we have $m = \mu\lambda$. In a length of string one wavelength long, therefore, the combined energy contributed by all the segments of length dx is

$$E_\lambda = \tfrac{1}{2}(\mu\lambda)\omega^2A^2. \tag{16.41}$$

The average power that must be supplied is thus

$$P_{av} = \tfrac{1}{2}\mu\lambda A^2\omega^2/T = \tfrac{1}{2}\mu A^2\omega^2 c, \tag{16.42}$$

where we have used the fact that $\lambda/T = c$ (Eq. 16.9). This result makes sense: It takes more effort to generate a wave of greater (angular) frequency or amplitude (in each case you must shake the end with a greater speed). Furthermore, as we would expect, a string of greater linear mass density requires

QUANTITATIVE TOOLS

more energy. Finally, if the wave speed increases, the energy is carried away more quickly, and so the rate at which it needs to be supplied increases.

EXAMPLE 16.5 Delivering energy

A wire with linear mass density $\mu = 0.0500$ kg/m is held taut with a tension of 100 N. At what rate must energy be supplied to the wire to generate a traveling harmonic wave that has a frequency of 500 Hz and an amplitude of 5.00 mm?

❶ **GETTING STARTED** The rate at which energy must be supplied to the wire is the power. Because none of the quantities given varies with time, the (instantaneous) power is equal to the average power, which is given by Eq. 16.42.

❷ **DEVISE PLAN** To calculate the average power, I need to know μ and A, both of which are given, as well as ω and c. The angular frequency ω is related to the frequency f by Eq. 15.4, $\omega = 2\pi f$, and I can obtain the wave speed c from Eq. 16.30.

❸ **EXECUTE PLAN** From Eq. 16.30 I obtain $c = \sqrt{\mathcal{T}/\mu} = \sqrt{(100\text{ N})/(0.0500\text{ kg/m})} = 44.7$ m/s. Equation 15.4 yields $\omega = 2\pi f = 2\pi(500\text{ Hz}) = 3.14\times10^3\text{ s}^{-1}$. Substituting these values into Eq. 16.42 gives

$$P_{av} = \tfrac{1}{2}(0.0500\text{ kg/m})(0.00500\text{ m})^2 \times (3.14\times10^3\text{ s}^{-1})^2(44.7\text{ m/s}) = 275\text{ W}. \checkmark$$

❹ **EVALUATE RESULT** The answer I obtain is a fairly large power—comparable to the power delivered by a 250-W light bulb or by a person exercising. The wave travels very fast, though, and the frequency is high, so the answer is not unreasonable.



16.21 Suppose that instead of shaking the end of a very long, taut string, you shake a point in the center of the string, keeping the amplitude and the frequency the same. Is the rate at which you must supply energy smaller than, equal to, or greater than if you shake the end?

16.9 The wave equation

In Chapter 15 I argued that simple harmonic oscillations are common because linear restoring forces are common. Linear restoring forces give rise to the equation of motion for a simple harmonic oscillator, which in turn gives rise to sinusoidally oscillating solutions. Waves are equally common because whenever a medium is disturbed it tends to oscillate sinusoidally and the oscillation couples to neighboring regions of the medium. As we shall show in this section, systems of coupled simple harmonic oscillators give rise to an equation called the wave equation, whose solutions are traveling waves. Here we derive this equation for the special case of a transverse wave traveling along a string; we shall encounter the same equation again when discussing electromagnetic waves.

Figure 16.42*a* shows a piece of a string that has been displaced from equilibrium by a passing wave. Let us focus on segment B between the positions x_i and $x_f = x_i + \Delta x$ and write the equation of motion for that segment. The segment is subject to two forces, one from each of the two adjacent pieces of string (we again ignore the force of gravity because it is much smaller than the tension). Figure 16.42*b* shows the free-body diagram for segment B. As we saw in Section 16.2, the x components of the forces are equal in magnitude to the tension $\mathcal{T}$ in the string and add to zero. The vector sum of the forces exerted on B is thus equal to the sum of the y components:

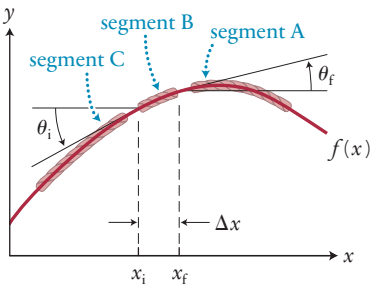
$$\Sigma F_{By} = F^c_{ABy} + F^c_{CBy} = \mathcal{T}\tan\theta_f - \mathcal{T}\tan\theta_i, \qquad (16.43)$$

where θ_i and θ_f are the directions of the tangents to the string at x_i and x_f. For small displacements, these angles are small, and so we can use the small-angle approximation $\tan\theta \approx \theta$ and rewrite Eq. 16.43 as

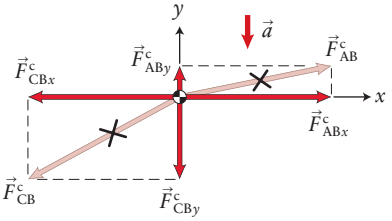
$$\Sigma F_{By} \approx \mathcal{T}(\theta_f - \theta_i) = \mathcal{T}\Delta\theta. \qquad (16.44)$$

Figure 16.42 Segment of a string transmitting a wave pulse.

(*a*) String transmitting wave pulse



(*b*) Free-body diagram for segment B



QUANTITATIVE TOOLS

The equation of motion for segment B is

$$\Sigma F_{By} = m_B a_{By} = (\mu \Delta x) a_{By}, \tag{16.45}$$

where m_B is the mass of segment B, a_{By} is the y component of its acceleration, and μ is its linear mass density. Substituting $\mathcal{T}\Delta\theta$ from Eq. 16.44 into the left side of Eq. 16.45 and rearranging terms yields

$$\frac{\Delta\theta}{\Delta x} = \frac{\mu}{\mathcal{T}} a_{By}. \tag{16.46}$$

In the limit $\Delta x \to 0$, the left side of this expression yields the derivative of the angle θ with respect to x. The angle θ is determined by the function f that describes the shape of the string in Figure 16.42*a*—the derivative of f with respect to x is equal to the slope, $\tan\theta$. A complication is that the shape of the string changes with time and so is a function of both x and t. We therefore have to use the *partial derivative* of $f(x, t)$ with respect to x, which means that we hold t constant while taking the derivative with respect to x (see Appendix B). The partial derivative of $f(x, t)$ with respect to x is written $\partial f/\partial x$, so we have

$$\tan\theta = \frac{\partial f}{\partial x}. \tag{16.47}$$

For small θ we can write $\tan\theta \approx \theta$, and so, in the limit $\Delta x \to 0$, the left side of Eq. 16.46 becomes the second partial derivative of $f(x, t)$ with respect to x:

$$\lim_{\Delta x \to 0} \frac{\Delta\theta}{\Delta x} \equiv \frac{\partial\theta}{\partial x} \approx \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \equiv \frac{\partial^2 f}{\partial x^2}. \tag{16.48}$$

Next we consider the right side of Eq. 16.46. The acceleration a_{By} is given by the second derivative with respect to time of the vertical position $y = f(x, t)$ of segment B. So the first partial derivative with respect to time $\partial f/\partial t$ gives the y component of the velocity of segment B, and the second partial derivative gives the y component of the acceleration:

$$a_{By} = \frac{\partial^2 f}{\partial t^2}. \tag{16.49}$$

Substituting Eqs. 16.48 and 16.49 into Eq. 16.46, we get

$$\frac{\partial^2 f}{\partial x^2} = \frac{\mu}{\mathcal{T}} \frac{\partial^2 f}{\partial t^2}, \tag{16.50}$$

and substituting Eq. 16.30, $c = \sqrt{\mathcal{T}/\mu}$, yields

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2}. \tag{16.51}$$

This is the one-dimensional **wave equation**. Any function of the form $f(x - ct)$ or $f(x + ct)$ is a solution to this equation (the proof of this is left as a homework exercise). All waves, mechanical and nonmechanical, result from equations of this form (but only for a taut string do we get Eq. 16.50 with $c = \sqrt{\mathcal{T}/\mu}$).

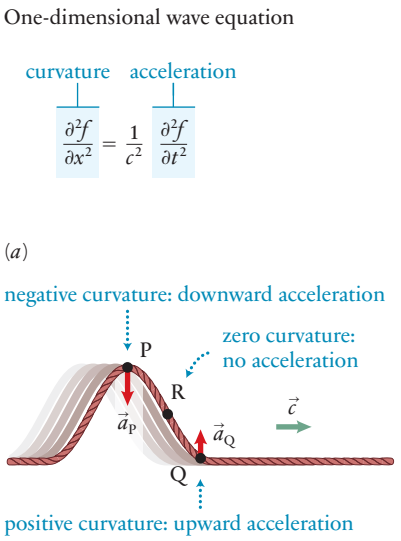
It is helpful to keep in mind what the terms in the wave equation tell us about the wave. The second partial derivative of f with respect to x on the left is related to the curvature of the string carrying the wave—the greater $\partial^2 f/\partial x^2$, the sharper the curvature of the string. The right side is the acceleration of particles of the string divided by the square of the wave speed. Thus the wave equation tells us that the curvature of the wave increases as the medium acceleration increases and decreases as the wave speed increases.

QUANTITATIVE TOOLS

For a fixed wave speed, the greater the acceleration of the particles in the medium, the greater the curvature. A downward curvature (concave down or $\partial^2 f/\partial x^2 < 0$) corresponds to a downward acceleration ($\partial^2 f/\partial t^2 < 0$). An upward curvature (concave up or $\partial^2 f/\partial x^2 > 0$) corresponds to an upward acceleration ($\partial^2 f/\partial t^2 > 0$). This is shown graphically in **Figure 16.43**. In Figure 16.43*a*, the curvature of the string is greatest (and downward) at the top. As we have seen, the acceleration is also greatest (and downward) at that point (compare with Figure 16.7*c*). At point Q, the curvature is upward; indeed, as the pulse travels to the right, point Q has to accelerate in the upward direction. At point R, the curvature and acceleration are both zero. Another illustration of the acceleration-curvature relationship at fixed wave speed is seen by comparing Figures 16.43*a* and *b*: If we move the end of the string up and down more quickly, the acceleration at point P increases and the pulse becomes narrower, giving a greater curvature at P.

The inverse relationship between curvature and wave speed is seen by comparing Figures 16.43*a* and *c*. With P (as well as all other points) having the same acceleration in the two drawings, the greater wave speed in part *c* stretches the string horizontally, decreasing the curvature.

Figure 16.43 Relationships among curvature, acceleration, and wave speed.



EXERCISE 16.6 Sinusoidal solution to the wave equation

Show that a sinusoidal traveling wave of the form $f(x, t) = A \sin(kx - \omega t)$ satisfies the wave equation for any value of k and ω.

SOLUTION The first partial derivative of $f(x, t)$ with respect to x is

$$\frac{\partial f}{\partial x} = kA \cos(kx - \omega t),$$

and so the second partial derivative with respect to x is

$$\frac{\partial^2 f}{\partial x^2} = -k^2 A \sin(kx - \omega t). \qquad (1)$$

The first partial derivative of $f(x, t)$ with respect to t is

$$\frac{\partial f}{\partial t} = -\omega A \cos(kx - \omega t).$$

Differentiating again gives

$$\frac{\partial^2 f}{\partial t^2} = -\omega^2 A \sin(kx - \omega t).$$

Multiplying both sides of this equation by $1/c^2$ and using $k = \omega/c$ (Eq. 16.11), I obtain

$$\frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = -\frac{\omega^2}{c^2} A \sin(kx - \omega t) = -k^2 A \sin(kx - \omega t). \qquad (2)$$

Because the right sides of Eqs. 1 and 2 are equal, the left sides must also be equal, so

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2},$$

which is Eq. 16.51, and so the wave equation is satisfied. ✔

16.22 If any time-dependent sinusoidal harmonic function, $f(x, t) = A \sin(kx - \omega t)$, satisfies the wave equation, what determines the values of the wave number k and the angular frequency ω for, say, a wave on a string?

Chapter Summary

Representing waves (Sections 16.1, 16.2, 16.5)

Concepts A **wave** is a *disturbance* that propagates through material (the medium) or through empty space.

A **wave pulse** is a single isolated propagating disturbance.

The **wave function** represents the shape of a wave at any given instant and changes with time as the wave travels.

The **wave speed** c is the speed at which a wave propagates. For a mechanical wave, c is different from the speed v of the particles of the medium and is determined by the properties of the medium.

The **displacement** $\vec{D}$ of any particle of a medium through which a mechanical wave travels is a vector that points from the equilibrium position of the particle to its actual position.

In a **transverse** mechanical wave, the particles of the medium move perpendicular to the direction of the pulse movement. In a **longitudinal** mechanical wave, these particles move parallel to the direction of the pulse movement.

In a **periodic wave,** the displacement at any location in the medium is a periodic function of time. A periodic wave is **harmonic** when the particle displacement can be represented by a sinusoidally varying function of space and time.

Quantitative Tools If a wave travels in the x direction with speed c and $f(x)$ describes the form (shape) of the wave, then the y component D_y of the displacement of a particle of the medium is

$$D_y = f(x - ct) \tag{16.3}$$

if the wave travels in the positive x direction and

$$D_y = f(x + ct) \tag{16.4}$$

if the wave travels in the negative x direction.

The **wavelength** λ of a periodic wave is the minimum distance over which the wave repeats itself.

For a periodic wave of period T, frequency f, and speed c, the **wave number** k is

$$k = \frac{2\pi}{\lambda}. \tag{16.7}$$

The wavelength λ is

$$\lambda = cT, \tag{16.9}$$

the angular frequency ω is

$$\omega = \frac{2\pi}{T}, \tag{16.11}$$

and the wave speed is

$$c = \lambda f. \tag{16.10}$$

For a transverse harmonic wave of amplitude A and initial phase ϕ_i traveling in the positive x direction, the y component D_y of the displacement of a particle of the medium is

$$D_y = f(x, t) = A \sin(kx - \omega t + \phi_i). \tag{16.16}$$

Combining waves (Sections 16.3, 16.4, 16.6)

Concepts **Superposition of waves:** The resultant displacement of two or more overlapping waves is the algebraic sum of the displacements of the individual waves.

Interference occurs when two waves overlap. The interference is *constructive* when the displacements due to the two waves are in the same direction and *destructive* when the displacements are in opposite directions.

If the displacement at a point in space remains zero as a wave travels through, that point is a **node**. The displacement at other points typically varies with time. If the displacement at a point in space varies over the greatest range as a wave travels through, that point is an **antinode**.

Quantitative Tools If two harmonic waves of equal wavelength λ and equal amplitude A travel in opposite directions on a string, they produce a standing wave. The y component D_y of the string displacement at any position x along the string is given by

$$D_y = 2A \sin kx \cos \omega t. \tag{16.20}$$

The nodes occur at

$$x = 0, \pm \frac{\lambda}{2}, \pm \lambda, \pm \frac{3\lambda}{2}, \ldots \tag{16.22}$$

and the antinodes occur at

$$x = \pm \frac{\lambda}{4}, \pm \frac{3\lambda}{4}, \pm \frac{5\lambda}{4}, \ldots. \tag{16.24}$$

When a wave pulse (the *incident* wave) reaches a boundary where the transmitting medium ends, the pulse is *reflected*, which means it reverses its direction.

When a wave pulse is reflected from a fixed boundary, the reflected pulse is *inverted* relative to the incident pulse. When the reflection is from a boundary that is free to move, the reflected pulse is not inverted.

A **standing wave** is a pulsating stationary pattern caused by the interference of harmonic waves of equal amplitude and wavelength traveling in opposite directions.

Waves on a string (Sections 16.7–16.9)

Quantitative Tools For a uniform string of mass m and length ℓ, the **linear mass density** μ (mass per unit length) is

$$\mu = \frac{m}{\ell}. \tag{16.25}$$

The speed of a wave on a string under tension $\mathcal{T}$ is

$$c = \sqrt{\frac{\mathcal{T}}{\mu}}. \tag{16.30}$$

The average power P_{av} that must be supplied to generate a wave of period T is

$$P_{\text{av}} = \tfrac{1}{2}\mu\lambda A^2\omega^2/T = \tfrac{1}{2}\mu A^2\omega^2 c. \tag{16.42}$$

Any function f of the form $f(x - ct)$ or $f(x + ct)$ that represents a wave traveling with speed c is a solution of the **wave equation:**

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}. \tag{16.51}$$

Questions and Problems

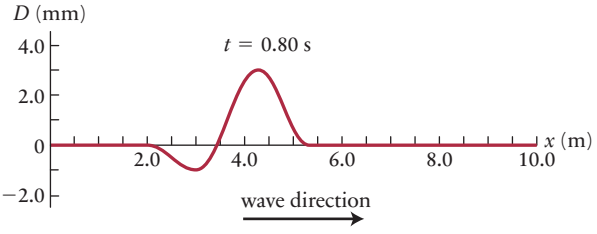
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

16.1 Representing waves graphically

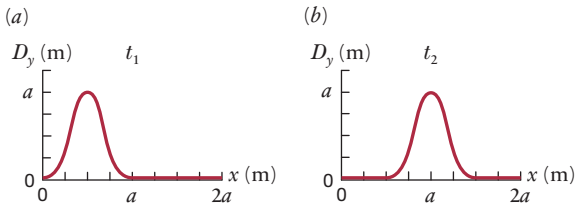
1. Is it possible for the same medium to carry both a longitudinal and a transverse wave? If so, give an example. If not, explain why not. •
2. Suppose you have a light spring stretched out and one end is attached to a wall. With this setup, you can move the free end in any of three directions (x, y, z). If the spring lies along, say, the z axis, which type of wave—transverse or longitudinal—do you create when you move the free end along the z axis? Along the x axis? Along the y axis? •
3. Figure P16.3 is a snapshot at $t = 0.80$ s of a wave pulse traveling on a string. Construct (*a*) the wave function at $t = 1.1$ s and (*b*) the displacement curve for the position $x = 3.0$ m. Assume that the wave pulse was created at the origin at $t = 0$. ••

Figure P16.3



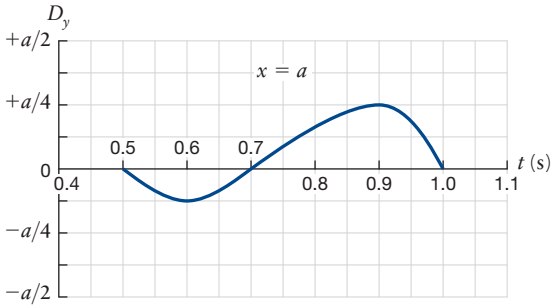
4. The graphs in Figure P16.4 show the displacement caused by a wave moving along a string at two instants, (*a*) t_1 and (*b*) t_2. Let v_{av} denote the average speed of a piece of string during the time interval between t_1 and t_2. Compare the wave speed c to v_{av} of a piece of the string at the position $x = a$. ●●

Figure P16.4



5. Figure P16.5 shows the displacement curve for the particle located at $x = a$ as a wave moves to the right along a string. If the wave advances a distance a each second, draw the wave function at the instant $t = 0$. ●●●

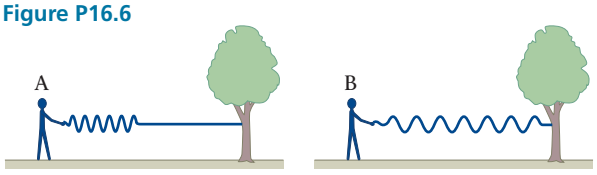
Figure P16.5



16.2 Wave propagation

6. Identical ropes were tied to two trees, and two men, A and B, started shaking the free ends at the same instant a short while ago (Figure P16.6). Which rope has the greater tension? Which man is shaking with the greater frequency? ●

Figure P16.6

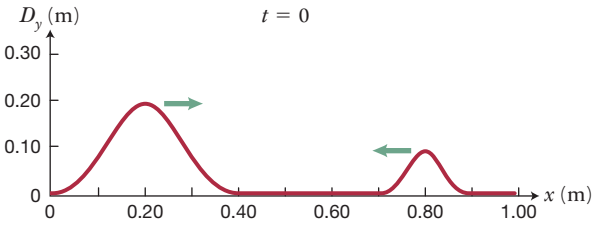


7. Waves propagating in a rope are observed to travel at 35 cm/s and their adjacent crests are 7.0 cm apart. Find the frequency and the period of the waves? ●
8. You and a friend each have one rope. You tie the two ropes together and stand as far apart as possible, each holding one end of the new longer rope and pulling to put it under tension. You then begin moving your arm in such a way as to produce a harmonic wave with a wavelength of 1.0 m. Your friend looks at the waves as they reach her arm. Is it possible that she measures a wavelength of (*a*) 0.8 m, (*b*) 1.0 m, or (*c*) 1.2 m? ●●
9. A ball of mass m_1 is attached on a rope and hangs vertically. The mass is slightly disturbed, and it generated a wave traveling through the rope. Suppose the ball is replaced with a heavier mass m_2 ($m_2 = 2m_1$) and the same disturbance is struck to the ball, compare the speed of the generated waves in both scenarios. ●●
10. A harmonic wave is made to travel along a string when you move your hand up and down. The wave has a specific period T_1, wavelength λ_1, amplitude A_1, and speed c_1, and also causes a certain transverse speed $v(x, t)$ of the particles that make up the rope. If you repeat this up-and-down motion, this time completing the same motion twice as fast as before, what happens to the (*a*) period, (*b*) wavelength, (*c*) amplitude, (*d*) speed, and (*e*) transverse speed? ●●●

16.3 Superposition of waves

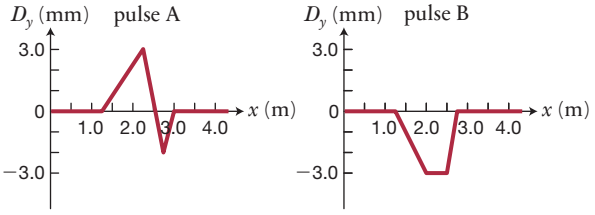
11. Water parks often have a pool with a wave-making machine at one end, consisting of a piston that pushes water back and forth. There are some places in the pool where the waves make inner-tube riders bounce up and down wildly, but other places where the water hardly moves. Explain what is happening. ●
12. Sinusoidal water waves are generated by a 18-W mechanical motor in an aquarium tank. How much the kinetic energy is in each 1.25-Hz wave generated? ●●
13. Figure P16.13 shows two waves in a rope approaching each other at instant $t = 0$. For any instant after $t = 0$, determine the maximum displacement that occurs (*a*) at $x = 0$, (*b*) at $x = 0.90$ m, and (*c*) anywhere along the rope. ●●

Figure P16.13



14. The two wave pulses in Figure P16.14 are traveling on the same string at $t = 1.0$ s. Sketch the shape of the string at this instant. ●●

Figure P16.14



15. Suppose two waves, identical except for the direction of travel, approach each other in a medium that obeys Hooke's law. At certain instants, the waves interfere constructively at all positions. At one such instant, when their peaks align, determine the kinetic energy of the wave function that results from the superposition of both waves. •••

16.4 Boundary effects

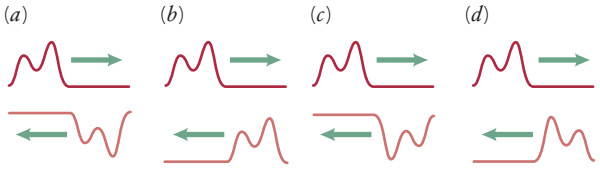
16. A harmonic wave traveling along a light string approaches a splice to a heavier string, as shown in Figure P16.16. Which changes as the wave crosses the boundary: wavelength, frequency, both, or neither? •

Figure P16.16



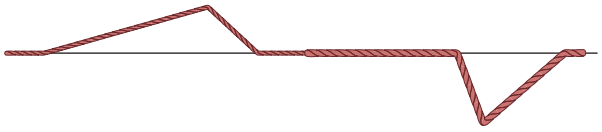
17. A transverse wave in a swimming pool reaches the concrete side and is reflected. Determine whether the reflected wave is inverted. •
18. A rope is attached to a tree trunk and made taut, and then a pulse traveling rightward is sent along the rope. For each case depicted in Figure P16.18, determine whether or not the reflected pulse is depicted correctly. ••

Figure P16.18



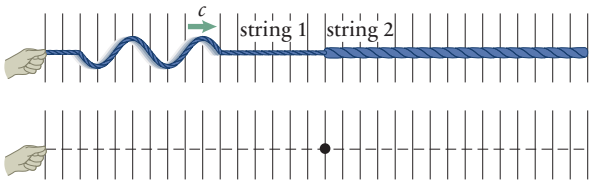
19. Figure P16.19 shows two strings of different linear mass densities, connected and held taut, at some instant. At some earlier instant, only one pulse existed on one of the strings. (*a*) Determine which string carried the initial pulse, and (*b*) sketch the approximate shape of the strings at the instant the initial pulse was created. ••

Figure P16.19



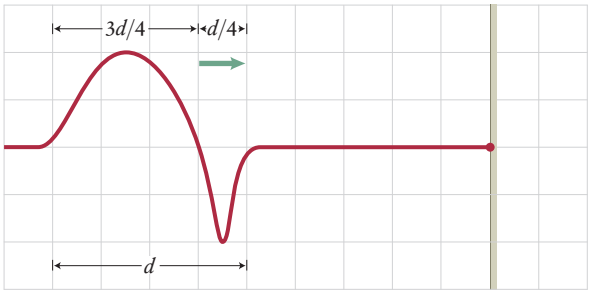
20. You shake the end of a taut string, creating two periods of a traveling sinusoidal wave, as shown in Figure P16.20. The string you are shaking (string 1) is connected to a second, much more massive, string (string 2). The wave speeds on the two strings differ by a factor of 2. (*a*) On the template, sketch the shape of the two strings a time interval $4T$ after the instant represented in Figure P16.20, with T the period of the original wave on string 1. Briefly explain how you determined the shape of the strings. (*b*) How do the periods of the waves on string 1 and string 2 compare? ••

Figure P16.20



21. Figure P16.21 shows a pulse as it approaches a fixed end of a rope. Draw the shape of the rope halfway through the reflection (that is, when half of the pulse length is still moving to the right and the other half of the pulse has already been reflected). •••

Figure P16.21



16.5 Wave functions

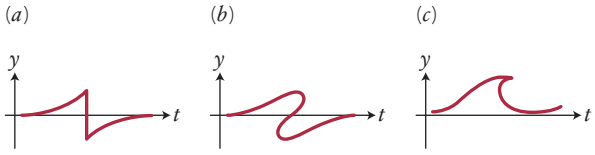
22. A kayaker on the near the shore observes 3 waves reaches the shore every 10 s. At what speed are the waves are moving if he estimates that adjacent crests are 5.0 m apart? •
23. While on a sailboat, you notice that the boat is moving up and down fifteen times periodically every minute. Find the period of oscillation and the distance between two adjacent wave crests if the waves are travelling at a rate of 4.5 m/s. ••
24. Participating in a human rally wave at a football game, you have to stand up every 15 s. (*a*) What is the wave frequency? (*b*) If the stadium is oval and the inner circumference of the oval is 0.56 km, what is the wave speed? Assume that the main wave pulse spans 30 people and that the distance between the heads of any two adjacent people is 1.0 m. ••
25. (*a*) Plot the time-independent wave function for the traveling pulse described by the time-dependent wave function

$$D_y(x, t) = \frac{a}{b^2 + (x - ct)^2},$$

with $a = 5.0\ \text{m}^3$, $b = 1.0$ m, and $c = 2.0$ m/s for $t = 0$, 1.0 s, 2.0 s, and 3.0 s. (*b*) What is the wave speed c of this pulse? ••

26. What is wrong with each displacement curve in Figure P16.26 ●●?

Figure P16.26



27. At $t = 0$, a wave pulse has a shape given by the time-independent wave function

$$f(x) = \frac{a}{b^2 + x^2},$$

where $a = 0.030\ \text{m}^3$ and $b = 2.0$ m. (*a*) If the pulse travels in the positive x direction at a wave speed of 1.75 m/s, write the time-dependent wave function $f(x, t)$ for this transverse wave. (*b*) Plot the time-dependent wave function $f(x, t)$ at $t = -0.50$ s and at $t = +0.50$ s over the range of x in which $f(x, t)$ is substantially changing. (*c*) Plot the displacement curve at $x = 2.0$ m. ●●

28. A wave traveling along a rope is described by

$$y(x, t) = (2.0\ \text{mm})\ \sin[(4.0\ \text{m}^{-1})x - (3.0\ \text{s}^{-1})t].$$

(*a*) Calculate the amplitude, wavelength, frequency, period, and speed of the wave. (*b*) Compute the y-component of the string from the equilibrium at $x = 0.30$ m and $t = 1.20$ s. ●●●

29. A certain wave is described by the function

$$f(x) = a\cos(b\pi x),$$

Where $a = 0.0750$ m and $b = 1.80\ \text{m}^{-1}$. (*a*) Calculate its wavelength. (*b*) If the wave is traveling at a speed of 12.5 m/s, what is its frequency f? (*c*) Find the angular frequency of the oscillation. ●●

30. As an earthquake starts, you are standing 150 km (as the wave travels) from the epicenter. A geophysicist near the epicenter immediately telephones you to let you know that the transverse wave from the earthquake is on its way to you and that it is described by the equation

$$f(x, t) = a\sin(bx - \omega t),$$

where $a = 0.560$ m, $b = 0.0157\ \text{m}^{-1}$, and a full wavelength passes any given location every 0.500 s. How long will it be before the ground you are standing on starts to shake? ●●

31. You are watching a large ripple tank and observed that the waves generated by a mechanical motor have a period of 3.5 s and the distance between two adjacent crests is 15 m. Write a possible general equation of the observed wave if the waves reach a vertical height of 0.30 m. ●●

32. A plucked violin string carries a traveling wave given by the equation

$$f(x, t) = a\sin[b(x - ct) + \phi_i],$$

with $a = 0.00580$ m, $b = 33.05\ \text{m}^{-1}$, and $c = 245$ m/s. (*a*) If $f(0, 0) = 0$, what is ϕ_i? (*b*) Determine the simple harmonic period of a point on the string. ●●

33. A wave on a rope is given by

$$f(x, t) = (0.750\ \text{cm})\cos\pi[(4.00\ \text{m}^{-1})x - (4.50\ \text{s}^{-1})t]$$

Calculate the (*a*) wavelength, (*b*) angular frequency, (*c*) frequency, (*d*) period, and (*e*) wave number. ●●●

34. The amplitude of a wave traveling on a string is 0.250 m. The 80.0-Hz wave is traveling in the positive x direction at a wave speed of 17.5 m/s. (*a*) Determine its wavelength, and write the equation for its time-independent wave function. (*b*) For $t = 3.00$ s, compute the displacement and transverse velocity of the point on the string at $x = 1.25$ m, assuming an initial phase ϕ_i of zero. (*c*) Write the time-dependent wave function for the wave. (*d*) If the string displacement at $x = 1.25$ m is 0.210 m at $t = 3.00$ s, write the time-dependent wave function that uses the smallest possible value of ϕ_i. ●●●

16.6 Standing waves

35. Watching your fish swim in their tank, you notice that when one fish repeatedly jumps, it causes a standing wave given by the time-independent function

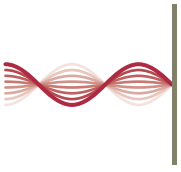
$$f(x) = a\sin(bx),$$

with $a = 0.015$ m and $b = 19.6\ \text{m}^{-1}$. If the tank is 0.96 m long, determine the positions of the first three nodes, beginning with the node at one end of the tank. ●

36. After tying one end of a rope to a stationary object, you flick the free end so as to cause a sinusoidal standing wave that has a maximum displacement of 0.50 m and a wavelength of 1.33 m. Write the time-dependent wave function in terms of the wave speed. ●

37. When one end of a string is tied to a pole and the other end is moved with frequency f, the standing wave pattern shown in Figure P16.37 is created. What is the smallest frequency at which the string can be moved to produce any standing wave? ●●

Figure P16.37



38. You have two string segments of equal length, segment 1 of linear mass density μ_1 and segment 2 of linear mass density $\mu_2 > \mu_1$. You splice the two together and then tie the combination tautly between two posts. You splice in such a way that, once attached to the posts, the length of

segment 1 to the left of the splice is the same as the length of segment 2 to the right of the splice. When you pluck at a particular frequency, you set up a standing wave that has this pattern: nodes at the splice and where each segment is attached to its pole plus two more nodes in one of the segments and three more nodes in the other segment (eight nodes in all). What is the numerical value of the ratio μ_1/μ_2? ●●

39. A guitar string has a length of 0.650 m and a mass of 4.00×10^{-3} kg. (*a*) If it is kept under a tension of 126 N, what is the fundamental (smallest harmonic) frequency at which it vibrates? (*b*) At what other frequencies (higher harmonics) could it vibrate at this fixed length? ●●

40. The equations for two waves traveling along the same string are

$$f_1(x, t) = a \sin(bx - qt)$$

and

$$f_2(x, t) = a \sin(bx + qt + \tfrac{1}{3}\pi),$$

with $a = 3.00 \times 10^{-2}$ m, $b = 4\pi$ m^{-1}, and $q = 500$ s^{-1}. (*a*) Calculate the amplitude and wavelength of the resultant displacement of the string at $t = 3.00$ s. (*b*) At $t = 1.70$ s, what is the displacement of the point on the string located at $x = 2.00$ m? ●●

41. A string vibrates in a standing wave according to the function

$$f(x, t) = a \sin(bx)\cos(qt),$$

with $a = 6.00 \times 10^{-2}$ m, $b = \frac{1}{3}\pi$ cm^{-1}, and $q = 40\pi$ s^{-1}. What are the amplitude and wave speed of each of the component waves that form this standing wave? ●●

42. A traveling wave on a long string is described by the time-dependent wave function

$$f(x, t) = a \sin(bx - qt),$$

with $a = 6.00 \times 10^{-2}$ m, $b = 5\pi$ m^{-1}, and $q = 314$ s^{-1}. You want a traveling wave of this frequency and wavelength but with amplitude 0.0400 m. Write the time-dependent wave function for a second traveling wave that could be added to the same string in order to achieve this. ●●

43. You have a trough 3.60 m long. It is filled with water, and you want to create a standing wave such that a node occurs every 0.30 m. (*a*) What is the wavelength of this wave? (*b*) A leaf floating on the water 0.60 m from one end of the trough is displaced a vertical distance $D_y = 10$ mm at $t = 3.2$ s, and the leaf's displacement has a period of 0.60 s. What is the resultant displacement of the standing wave? (*c*) Write the time-dependent wave function for the standing wave. ●●

44. Two sinusoidal waves travel along the same string. Their time-dependent wave functions are

$$f_1(x, t) = a \sin(bx - qt - \tfrac{1}{4}\pi)$$

$$f_2(x, t) = a \sin(bx - qt - \tfrac{1}{3}\pi),$$

with $a = 5.00 \times 10^{-2}$ m, $b = 0.120$ m^{-1}, and $q = 180$ s^{-1}. (*a*) What is the phase difference between these two waves? (*b*) Write the time-dependent wave function for the wave created by the superposition of these two waves. (*c*) What is the displacement of the string at $x = 2.00$ m at $t = 1.70$ s? ●●●

45. Two sinusoidal waves travel in opposite directions along the same string. The wavelength and frequency are the same in both waves, and each has amplitude 0.0289 m. If there is a phase difference $\Delta\phi$ between the two waves, what is the maximum displacement of the wave created by the superposition? ●●●

16.7 Wave speed

46. A piano wire has a linear mass density of 5.00×10^{-3} kg/m and is under 1.35 kN of tension. What is the frequency of the wave if it has a wavelength of 20 cm? ●●

47. What must be the tension applied to a 12-m plastic string weighing 43.5 N to produce a wave that propagates at a speed of 25 m/s? ●●●

48. A large crane is being used to lower a 10,000-kg storage unit into place on a barge. If the cable that supports the unit from the end of the crane is 12 m long, and the linear mass density of the cable is 1.78 kg/m, how long does it take a wave to travel the length of the cable? ●

49. A violinist is bowing one string to produce a certain note. List three ways in which she can produce a note of higher frequency, either while playing or by preparing the violin differently before she begins to play. ●●

50. How long after the propellers at the back of an oil tanker start turning does the front of the ship start to move? ●●

51. A wave traveling along a string has a speed of 24 m/s when the tension in the string is 120 N. What is the speed of the same wave if the tension is reduced to 100 N? ●●

52. Two wires that have different linear mass densities, 0.45 kg/m and 0.29 kg/m, are spliced together. They are then used as a guy line to secure a telephone pole. If the tension is 300 N, what is the difference in the speed of a wave traveling from one wire to the other? ●●

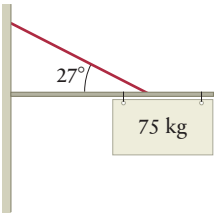
53. A string of linear mass density μ_1 fixed at both ends is replaced by a string of linear mass density $16\mu_1$. By what factor does the frequency of the fundamental standing wave change if the tension and wavelength remain what they were in the original string? ●●

54. A 25.0-m steel wire and a 50.0-m copper wire are attached end to end and stretched to a tension of 145 N. Both wires have a radius of 0.450 mm, and the densities are 7.86×10^3 kg/m^3 for the steel and 8.92×10^3 kg/m^3 for the copper. (Note that these are mass densities, mass per unit volume, not linear mass densities, mass per unit length.) How long does a wave take to travel from one end to the other end of the combination wire? ●●

55. A 100-m steel cable that helps support the Golden Gate Bridge is 72.0 mm in diameter and composed of 100 steel wires twisted together. Approximate this as a single uniform steel cable that has a mass density of 7.86×10^3 kg/m^3. The wind generates in the cable a vibration that creates a wave that travels down the cable at 380 m/s. What is the tension in the cable? ●●

56. One end of the wire in Figure P16.56 is anchored to the side of a building, and the other end is attached at an angle of 27° to a light horizontal bar from which a 75-kg sign hangs. The linear mass density of the wire is 0.067 kg/m. If the wind causes the sign to vibrate so that a wave travels along the wire, at what speed does the wave travel? •••

Figure P16.56



16.8 Energy transport in waves

57. A taut rope has a mass of 0.128 kg and a length of 3.60 m. What average power must be supplied to the rope to generate sinusoidal waves that have amplitude 0.200 m and wavelength 0.600 m if the waves are to travel at 25.0 m/s? •
58. Consider two waves X and Y traveling in the same medium. The two carry the same amount of energy per unit time, but X has half the amplitude of Y. What is the ratio of their wavelengths? Which wave has the longer wavelength? •
59. Transverse waves are being generated on a rope being held under constant tension. By what factor is the power required to generate the wave increased or decreased (*a*) if the length of the rope is halved, (*b*) if the amplitude and angular frequency are doubled, and (*c*) if the wavelength and rope length are doubled? ••
60. You hold one end of a string that is attached to a wall by its other end. The string has a linear mass density of 0.067 kg/m. You raise your end briskly at 12 m/s for 0.016 s, creating a transverse wave that moves at 31 m/s. (*a*) How much work did you do on the string? What are the wave's (*b*) energy, (*c*) potential energy, and (*d*) kinetic energy? ••
61. One end of a horizontal string that has a linear mass density of 3.5 kg/m is displaced vertically at a speed of 45 m/s for 6.7 ms. The pulse created by this motion travels down the string at 78 m/s. How much kinetic energy travels along the string? ••
62. A wave traveling along a string is described by the time-dependent wave function $f(x, t) = a \sin(bx + qt)$, with $a = 0.0268$ m, $b = 5.85\ \text{m}^{-1}$, and $q = 76.3\ \text{s}^{-1}$. The linear mass density of the string is 0.0456 kg/m. Calculate (*a*) the wave speed c, (*b*) the wave frequency f, and (*c*) the power P supplied by the wave. ••
63. An oscillator attached to a string transmits 350 mW of power down the string. The wave created by this motion has amplitude 0.0165 m and wavelength 1.80 m. If the string is under 14.8 N of tension, what is the wave speed? ••
64. A cord of linear mass density 0.360 kg/m is attached to a harmonic oscillator that has a maximum output of 200 W. (*a*) With the oscillator operating at this maximum power and the tension in the cord set at 30.0 N, a wave of amplitude 8.00×10^{-3} m is created. What is the minimum wavelength possible for a wave of this amplitude? (*b*) What is the power output of the oscillator if the wave generated has the same amplitude but double the wavelength you calculated in part *a*, and the tension is 15.0 N? •••
65. One end of a 9.00-m wire vibrates sinusoidally, creating a wave that travels in the positive x direction. The wave has a frequency of 60.0 Hz, a wavelength of 3.00 m, and an amplitude of 0.0725 m. (*a*) Write the time-dependent wave function that describes this wave. (*b*) If the mass of the wire is 1.20×10^{-2} kg, what is the power supplied to the wire to create this wave? •••

16.9 The wave equation

66. Show that $f(x, t) = e^{b(x-vt)}$, where b and v are constants, is a solution to Eq. 16.51, the wave equation. •
67. What keeps a surfer riding the crest of a wave from dropping down to the wave's base? ••
68. A string initially hanging vertically is displaced to the right with a speed of 25 m/s for 0.040 s, then returned to its original position in 0.010 s. This motion is repeated, with no delays, numerous times. The wave created by this motion travels down the string at 67 m/s and has a period of 0.050 s. (*a*) Write the time-dependent wave function for this pulse, and (*b*) show that this function is a solution to the wave equation (Eq. 16.51). ••
69. Show that if $f_1(x, t) = A_1 \sin(k_1 x - \omega_1 t)$ and $f_2(x, t) = A_2 \sin(k_2 x - \omega_2 t)$ are both solutions to the wave equation (Eq. 16.51), then $f(x, t) = f_1(x, t) + f_2(x, t)$ is also a solution. ••
70. (*a*) Show that the equations (*i*) $A \cos(kx + \omega t)$, (*ii*) $e^{-b|x-qt|^2}$, and (*iii*) $-(b^2 t - x)^2$ all satisfy the wave equation (Eq. 16.51) provided that A, k, ω, b, and q are constants. (*b*) Calculate the speed of each wave of part *a*. •••

Additional Problems

71. Two children are holding a rope taut between them when one suddenly yanks the rope. (*a*) Does the child at the other end instantly get pulled? (*b*) Answer the same question for two children holding a rigid steel bar between them. •
72. A heavy rope is hanging by one end from the ceiling with a block hanging from the other end. Is the speed of a pulse in the rope greater at the top or the bottom of the rope? •
73. First you stretch a spring to twice its unstretched length and send pulse A from one end to the other. Then you stretch the spring to three times its unstretched length and send pulse B from one end to the other. How does the time interval needed for A to travel through the spring compare with the time interval needed for B to do so? ••
74. You pluck a guitar string so as to produce a standing wave for which $\lambda = 1.3$ m. What is the maximum value of the string's displacement at a position 0.500 m from the lower end of the string if the maximum resultant displacement of the wave is 0.75 mm? ••

75. Tossing rocks into the center of a pond, you notice that the ripples are evenly spaced 170 mm apart. If it takes 5.0 s for 15 ripples to pass a log sticking up out of the water, (*a*) what is the wave speed? (*b*) How long does it take a ripple to reach shore if the pond is 12 m in diameter? ●●

76. When a person sings in the shower, certain notes seem louder than others. What's going on? ●●

77. A standing wave is created on a string that is 1.75 m long. The wave has eight nodes counting the nodes at the two ends. (*a*) What is the wavelength? (*b*) What is the frequency if the wave speed is 130 m/s? (*c*) If the resultant displacement is 20.0 mm, write the time-dependent wave function that describes the standing wave, including units where appropriate. ●●

78. Would water waves traveling along a trough on the surface of Mars travel faster than, slower than, or at the same speed as water waves traveling along a similar trough on the surface of Earth? (Hint: You'll need data from Table 13.1 on page 430.) ●●

79. You hold one end of a horizontal string and raise your hand during a time interval of 0.18 s, doing 1.2 J of work on the string in the process. You then allow your hand to return to its original position. If the wave pulse you create travels along the string at 67 m/s and the linear mass density of the string is 0.0282 kg/m, what was the vertical speed of your hand? ●●

80. Your friend wants to test your understanding of waves. She weaves together one end of two different ropes (one with greater mass density than the other, but each with the same radius and length) to make a single rope. Then she attaches one end of the combination to a pole. You are blindfolded and asked to hold the other end of the stretched rope. Your friend expects you to determine not only which end you are holding (the heavier or lighter rope), but also whether your friend used a fixed or free method of attachment to the pole. You think for a minute, then laugh and declare that you can even estimate the ratio of the linear mass densities! While she is speechless, you create a single wave pulse with your hand. ●●● CR

81. Your job as assistant coach is to attach one end of several heavy, 10-m ropes to the trusses that support the gym ceiling, leaving the other end of each hanging free. The team is going to spend the afternoon climbing these ropes, but you begin to wonder about wave pulses traveling on such a rope. You could shake the free end of a rope back and forth once and measure how long it takes the pulse created by this motion to arrive at the top of the rope, but you are more of a calculation person. ●●● CR



17

Waves in Two and Three Dimensions

CONCEPTS

17.1 Wavefronts
17.2 Sound
17.3 Interference
17.4 Diffraction

QUANTITATIVE TOOLS

17.5 Intensity
17.6 Beats
17.7 Doppler effect
17.8 Shock waves

Now that we have treated waves that propagate in one dimension, we can extend our description to waves that propagate in more than one dimension—for example, waves on a water surface, which propagate in two dimensions, and sound waves, which propagate in three dimensions. The basic mechanism for these waves is the same as that for one-dimensional waves: When particles in a medium are coupled together, a disturbance caused at one point in a medium propagates to the surrounding points. Therefore, most of what we know about waves in one dimension carries over to the waves treated in this chapter. However, there are important differences between waves in one dimension and waves in more than one dimension. For example, the amplitude of waves propagating in more than one dimension decreases even in the absence of energy dissipation, and wave interference in more than one dimension gives rise to a number of phenomena not observed in one dimension.

17.1 Wavefronts

When you drop a stone in water, the stone disturbs the water surface and causes a number of circular ripples to travel outward along the water surface, emanating from the point at which the stone entered the water. The radius of each ripple starts out small and increases as the ripple moves outward at constant speed. The ripples constitute a two-dimensional **surface wave,** which is a wave that propagates in two dimensions. Surface waves are analogous to the one-dimensional waves we studied in Chapter 16, where propagation is always along a line.

If, instead of dropping a stone, you move a stick sinusoidally up and down in the water, you generate a periodic surface wave. **Figure 17.1** shows two cutaway views of a periodic surface wave at two instants that are half a period apart. The graph below each view shows the vertical component of the displacement of the water surface as a function of the radial distance r to the source—in other words, the *wave function* of the wave. Because the wave is circular, the wave function is the same in all radial directions. Note the similarity between the shape of the curves in these graphs and the shape of the wave in Figure 16.9 for a periodic wave on a string. Because the surface wave has circular symmetry, all points that are equidistant from the source have the same phase of oscillation. For example, at instant t_i all points that are a distance λ from the source have reached their maximum displacement above the equilibrium surface, forming a circular crest. Half a period later, the same points have reached their maximum displacement below the surface, forming a circular trough.

When the source of the wave can be localized to a single point in space—a small stone dropped in a pond, the stick in Figure 17.1—the source is said to be a **point source.** Point sources are treated as if they have no physical extent. The actual source need not be physically small, however. Stars, for example, are much larger than Earth, yet because they are so far away they can be treated as a point source of light (a wave). Likewise, in Figure 17.1, if the diameter of the stick is small compared to the diameter of the ripples, we can consider the stick to be a point source.

The blue region of **Figure 17.2** shows a top view of the circular crests and troughs of a periodic surface wave spreading out from a point source. We can schematically represent the wave by a series of circles

Figure 17.1 Cutaway views of a wave on the surface of a liquid and the corresponding vertical displacement of the surface as a function of distance r from the oscillating wave source, which is located at $r = 0$.

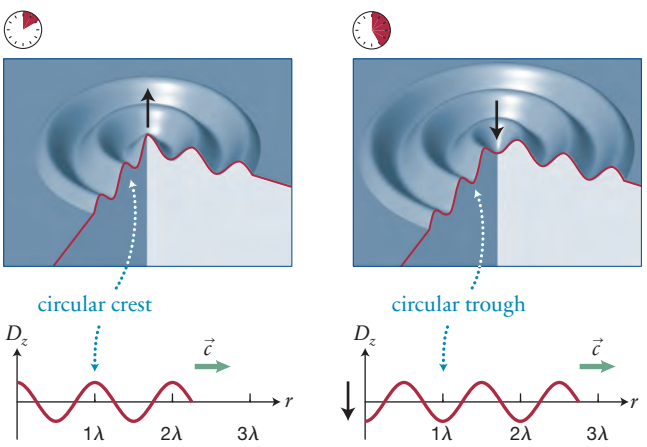
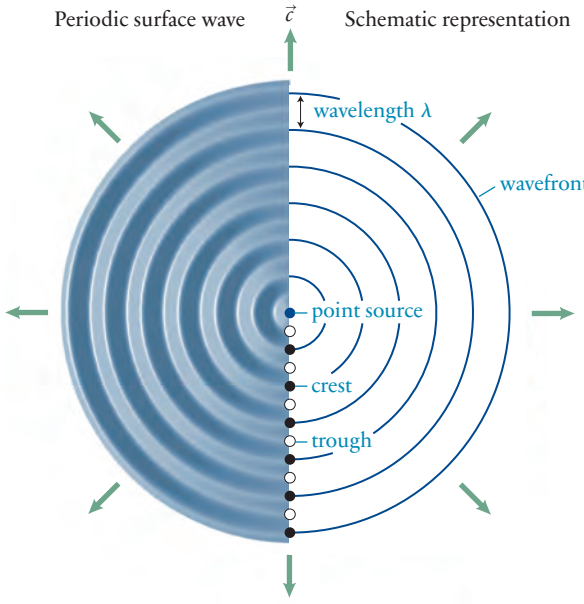


Figure 17.2 Top view of the circular crests and troughs formed in a periodic surface wave. We represent the wave schematically as a series of wavefronts.



CONCEPTS

Figure 17.3 A surface wave moving at speed c. Because surface waves expand in two dimensions from the point source (at $t = 0$), the wave energy is distributed over a larger and larger circumference.

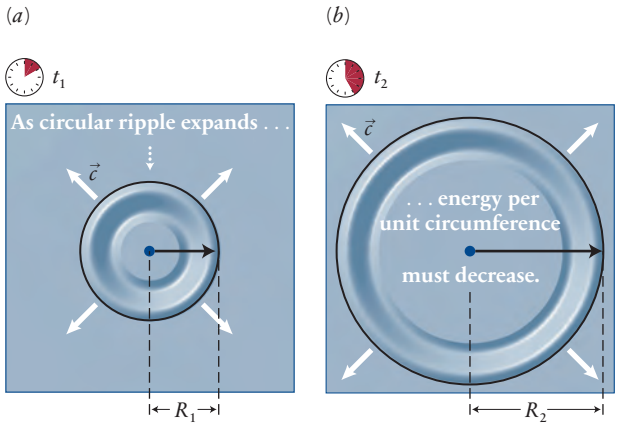
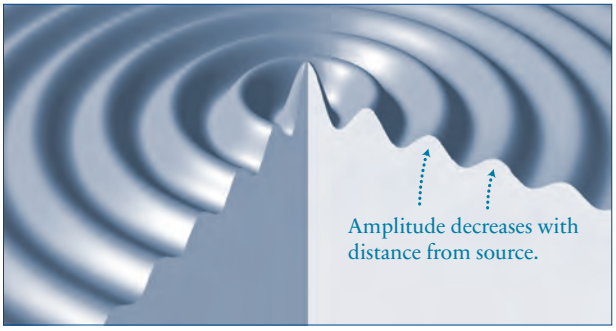


Figure 17.4 Cutaway view of a periodic surface wave. Because the energy in a given wavefront is spread out over an increasingly large circumference as the wavefront moves away from the source, the amplitude decreases as the wavefront moves away from the source.



separated by the wavelength λ (white region). Each circle corresponds to points on the surface that have the same phase of oscillation. Curves or surfaces on which all points have the same phase constitute a **wavefront.** As the wave travels away from the source at the wave speed c, the radius of each wavefront increases.

This expansion of circular surface waves causes the wave energy to be spread out over a larger and larger region. Consider, for example, the spreading ripples in **Figure 17.3**. At instant t_1 the ripples cross a circle of radius R_1 centered on the wave source; at instant t_2 they are about to cross a circle of radius R_2. If there is no loss of wave energy as the wave moves from R_1 to R_2—that is, if there is no dissipation of energy—the amount of energy that crosses the circle of radius R_2 must be the same as the amount that crosses the circle of radius R_1. Because the circumference increases, there is less energy per unit length along the circumference at t_2 than there is at t_1.

CONCEPTS

17.1 Let $t_2 = 2t_1$ in Figure 17.3. (*a*) How does R_1 compare with R_2? (*b*) If the energy in the wave is E and there is no dissipation of energy, what is the energy per unit length along the circumference at R_1? At R_2? (*c*) How does the energy per unit length along a wavefront vary with radial distance r?

As Checkpoint 17.1 shows, the expansion of circular wavefronts causes the energy per unit length along a wavefront to decrease as $1/r$. Therefore, it follows that the wave amplitude is proportional to $1/\sqrt{r}$. (You can see this amplitude decrease in the cutaway view in **Figure 17.4**.) Keep in mind that this amplitude decrease has nothing to do with energy dissipation: It is due entirely to the spreading out of the wave, which causes the energy to be distributed over a larger circumference.

Some waves—sound waves and light waves are two examples—spread out in three dimensions. If the source is a point source and the spreading is uniform in all three dimensions, the wavefronts are spherical (**Figure 17.5**). The resulting waves are called **spherical waves.** As a three-dimensional wave propagates, each wavefront expands outward at the wave speed c, and the energy carried by the wavefront is spread out over a larger and larger spherical area. When a wavefront reaches a distance r from the source, the energy in the wavefront is spread out over a surface area of $A = 4\pi r^2$, and so the energy per unit area carried by the wavefront is proportional to $1/r^2$ (**Figure 17.6**). The amplitude of a spherical wave is therefore proportional to $1/r$ (Eq. 16.41 again).

Because it is difficult to represent wavefronts in three dimensions, we usually represent spherical wavefronts by a series of circles, just as we do for two-dimensional waves. Keep in mind, however, that there

Figure 17.5 The wavefronts from a point source emitting waves in three dimensions are uniformly expanding, concentric spheres.

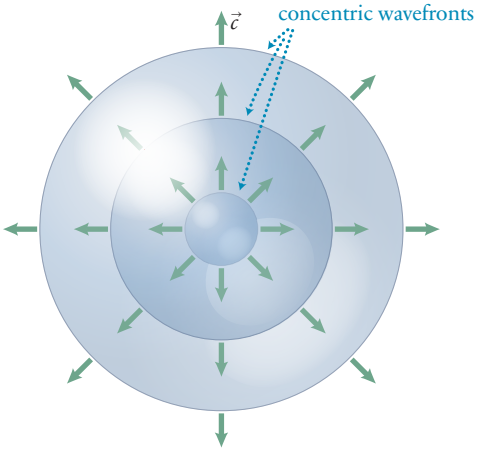
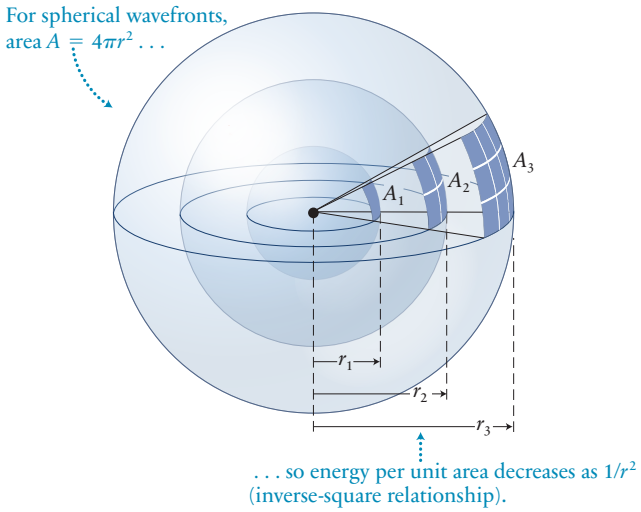


Figure 17.6 The inverse-square relationship between energy and radius for a wave expanding uniformly in three dimensions without dissipation. All the energy that crosses area A_1 also crosses A_2 and A_3. Because the area of a sphere is proportional to r^2 ($A_{\text{sphere}} = 4\pi r^2$), the energy per unit of surface area decreases as $1/r^2$.



is an important difference between waves in two and in three dimensions:

> **The amplitude of waves in two dimensions decreases with distance r from the source as $1/\sqrt{r}$; in three dimensions it decreases as $1/r$. This decrease is due purely to the spreading out of the wavefronts and involves no loss of energy.**

EXAMPLE 17.1 Ripple amplitude

The amplitude of a surface wave for which $\lambda = 0.050$ m is 5.0 mm at a distance of 1.0 m from a point source. What is the amplitude of the wave (*a*) 10 m from the source and (*b*) 100 m from the source? During a time interval equal to 100 periods, by what percentage does the amplitude of a wavefront decrease (*c*) after the wavefront passes a position 1.0 m from the source and (*d*) after it passes a position 100 m from the source?

1 GETTING STARTED I am given that the amplitude $A = 5.0$ mm at $r = 1.0$ m. As the wave spreads out, its amplitude diminishes, and I need to calculate the amplitude at $r = 10$ m and $r = 100$ m. In addition I need to determine by how much the wave is reduced as it propagates over a 100-period time interval past these two positions.

2 DEVISE PLAN Because the wave is a two-dimensional surface wave, the amplitude is proportional to $1/\sqrt{r}$. I know the amplitude $A_{1.0\text{ m}}$ at $r = 1.0$ m, so I can use this dependence to determine the amplitude at other distances from the source. For parts *a* and *b*, I need to determine $A_{10\text{ m}}$ and $A_{100\text{ m}}$ at $r = 10$ m and $r = 100$ m. In 100 periods, the wavefront travels a distance equal to 100 wavelengths, which is $100 \times (0.050\text{ m}) = 5.0$ m, so for parts *c* and *d*, I also need to determine the amplitudes 5 m beyond those positions.

3 EXECUTE PLAN (*a*) The ratio of the amplitudes at 1.0 m and 10 m is $\sqrt{(1.0\text{ m})}/\sqrt{(10\text{ m})} = \sqrt{(1.0/10)} = 0.32$, and so the amplitude at 10 m is $0.32 \times (5.0\text{ mm}) = 1.6$ mm. ✔

(*b*) At 100 m, $\sqrt{(1.0\text{ m})}/\sqrt{(100\text{ m})} = 0.10$, and so the amplitude is $0.10 \times (5.0\text{ mm}) = 0.50$ mm. ✔

(*c*) 100 periods after passing the 1.0-m position, the wavefront is 6.0 m from the source and the amplitude decreases by a factor of $\sqrt{(1.0\text{ m})}/\sqrt{(6.0\text{ m})} = 0.41$. This means $A_{6.0\text{ m}}$ is 41% of $A_{1.0\text{ m}}$: $A_{6.0\text{ m}} = 0.41(5.0\text{ mm}) = 2.0$ mm, a decrease of about 60%. ✔

(*d*) Then 100 periods after passing the 100-m position, the wavefront is 105 m from the source and the amplitude decreases by a factor of $\sqrt{(100\text{ m})}/\sqrt{(105\text{ m})} = \sqrt{(1/1.05)} = 0.98$. This means that $A_{105\text{ m}}$ is 98% of $A_{100\text{ m}}$: $A_{105\text{ m}} = 0.98(0.50\text{ mm}) = 0.49$ mm, a decrease of 2%—a mere 0.02% each period. ✔

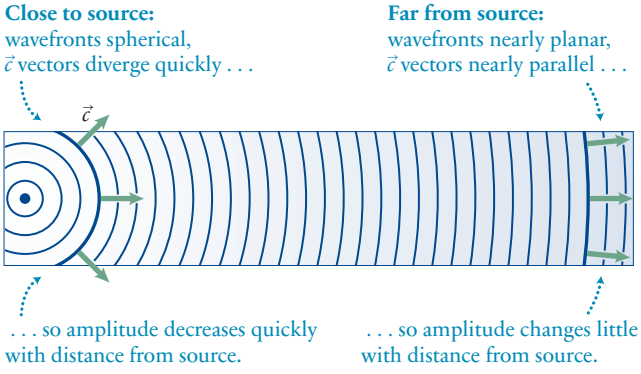
4 EVALUATE RESULT The amplitudes at 10 m and 100 m are both smaller than the amplitude at 1.0 m, which is what I expect. That the amplitude decrease over 100 wavelengths is less farther away from the source agrees with what I expect from the $1/\sqrt{r}$ dependence.

As this example demonstrates, the decrease in amplitude is most pronounced close to the source; far from the source, the decrease is hardly noticeable. **Figure 17.7** illustrates the same point in a different way: Far from the source, the wavefronts are nearly flat rather than curved, and as a wavefront straightens, a fixed length of it diverges less. Far from a point source, a spherical wavefront essentially becomes a two-dimensional flat wavefront called a **planar wavefront.** Likewise, far from a point source, a circular surface wavefront essentially becomes a one-dimensional straight wavefront.

17.2 Notice that in the views of the surface wave in Figure 17.1 the amplitude does not decrease with increasing radial distance r. How could such waves be generated?

CONCEPTS

Figure 17.7 Far from the source, wavefronts are nearly flat, so amplitude changes little with additional distance from the source.



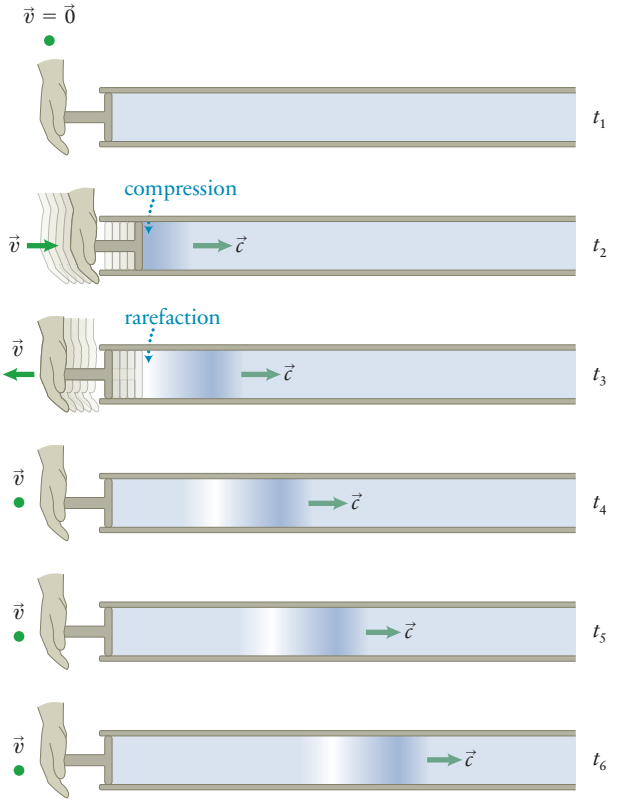
17.2 Sound

The human ear can detect some longitudinal waves at frequencies in the range from about 20 Hz to 20 kHz. Such waves, which propagate through any kind of material—solid, liquid, or gas—constitute what we call **sound.** Sound propagates through any medium that can exert elastic restoring forces. If you find it hard to think of a gas as being able to exert an elastic force, squeeze an empty plastic bottle first with a tightly sealed cap and then without it. Without the cap, squeezing the bottle simply pushes some air out of the bottle. With the cap, the bottle exerts an elastic force because of the air in it. It is this elastic force that is responsible for transmitting sound waves.

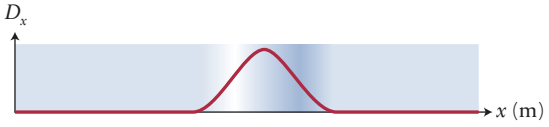
Figure 17.8 shows a simple wave pulse propagating through a long, air-filled tube. One end of the tube is fitted with a piston. Starting at t_2, the piston is rapidly moved back and forth, first pushing the molecules in the layer of air adjoining the piston closer together and then letting them expand again. These molecules, in turn, push on the molecules in a neighboring layer, and so on, creating a disturbance that propagates through the air in the tube. The disturbance consists of a *compression* (region where the molecules are crowded together) and a *rarefaction* (region where they are spaced far apart). The resulting wave is called a *compressional wave*. The wave speed c at which the disturbance travels through the tube depends on the density and the elastic properties of the medium. For dry air at a temperature of 20 °C, the speed of sound waves is 343 m/s.

Figure 17.8 A piston generates a longitudinal sound pulse that propagates along an air-filled tube. The disturbance consists of a compression and a rarefaction. The wave function is similar to the one for a longitudinal wave in a spring (see Figure 16.6).

(*a*) Longitudinal wave pulse created by in-and-out movement of piston



(*b*) Wave function of pulse at instant t_6



EXERCISE 17.2 Wavelength of audible sound

Given that the speed of sound waves in dry air is 343 m/s, determine the wavelengths at the lower and upper ends of the audible frequency range (20 Hz–20 kHz).

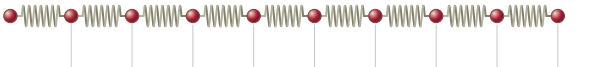
SOLUTION The wavelength is equal to the distance traveled in one period. At 20 Hz, the period is $1/(20\text{ Hz}) = 1/(20\text{ s}^{-1}) = 0.050$ s, so the wavelength is $(343\text{ m/s})(0.050\text{ s}) = 17$ m. ✔ The period of a wave of 20 kHz is $1/(20{,}000\text{ Hz}) = 5 \times 10^{-5}$ s, so the wavelength is $(343\text{ m/s})(5.0 \times 10^{-5}\text{ s}) = 17$ mm. ✔

The wave pulse in Figure 17.8 is longitudinal—the air molecules are displaced in the same direction as that in which the wave propagates. In bulk gases and liquids that have no internal friction, there can be no transverse waves because such gases and liquids do not exert elastic forces in the transverse direction. Notice the similarity between the wave pulse shown in Figure 17.8 and the longitudinal wave pulse propagating through a coil spring shown in Figure 16.5.

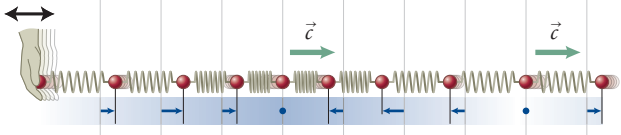
Figure 17.9a shows a chain of identical beads coupled together by springs, which can serve as a simple

Figure 17.9 A simple mechanical model for longitudinal waves.

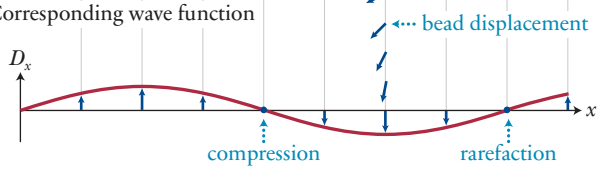
(*a*) Identical beads coupled by springs



(*b*) Snapshot of longitudinal wave propagating along chain



(*c*) Corresponding wave function



CONCEPTS

mechanical model for studying longitudinal waves. A snapshot of a harmonic longitudinal wave propagating through the chain is shown in Figure 17.9*b*. The leftmost bead is made to oscillate back and forth harmonically. Because each bead executes, in turn, the same longitudinal oscillation, a compressional wave propagates along the chain. Although the beads move in the same direction as the wave, there is no permanent displacement of beads in that direction. What are transported along the chain, however, are energy (as kinetic energy of the beads and elastic potential energy of the springs) and momentum.

17.3 Does the wave speed along the chain shown in Figure 17.9 increase or decrease when (*a*) the spring constant of the springs is increased and (*b*) the mass of the beads is increased?

From Section 16.1, you know that a snapshot of a transverse wave gives us the wave function directly. Things are not so simple with longitudinal waves, however. Because with these waves the displacement of the medium is along the axis of wave propagation, it is cumbersome to interpret any wave snapshots. The wave function shown in Figure 17.9*c*, however, provides an alternative representation of the wave and reveals its harmonic nature, which is not immediately obvious in Figure 17.9*b*. This graph is obtained by plotting the displacement of each bead from its equilibrium position versus the bead's equilibrium position. To create Figure 17.9*c*, the displacement vector of each bead in Figure 17.9*b* is rotated 90° counterclockwise and its tail placed on the x axis at the *equilibrium* position of the bead given in Figure 17.9*a*.

The perception of sound

The human ear is capable of detecting longitudinal waves transmitted through a gas, liquid, or solid medium in what is called the *audible frequency range*. This range, which varies with age and also from person to person, extends from about 20 Hz to 20 kHz. The sensitivity of the ear varies greatly over this range, however, peaking at about 3 kHz. Waves outside this range cause no auditory nerve signals in humans, although large-amplitude waves of very low frequency, called *subsonic waves* (such as earthquakes), can be felt. Some animals, such as bats, can detect longitudinal waves of frequencies up to 120 kHz, called *ultrasonic waves* because they are above the audible frequency range for humans.

The perception of sound is a complicated and subjective matter. The mechanical part of this process, which is determined by the physiology of the ear, is well understood. A sound wave that reaches the ear propagates through a complex system of solid and liquid parts of the outer and middle ear to the inner ear, where the wave causes *hair cells* to oscillate. These oscillations cause the auditory nerve to send signals to the brain.

The neurological processing of these signals by the brain is not yet fully understood. The properties of sound are typically described using terms such as *loudness*, *pitch*, and *timbre* or *tonal quality*. These terms are subjective, however, and so there is no simple correlation between them and the parameters used to describe waves in physics (amplitude, frequency, harmonic content).

Loudness: The perception of loudness is connected to sound wave energy and therefore amplitude. Because the ear is capable of detecting waves over a large range of wave energies—from the very faint sound of a pin drop to the loud roar of an airliner taking off—the response is highly nonlinear (discussed in Section 17.5). The correlation between loudness and energy is complicated by the fact that the perception of loudness also depends on frequency: A 1-kHz wave sounds louder than a 200-Hz wave of equal energy (see Figure 17.26).

Pitch: The sensation of pitch is related to the physical property of frequency. The correlation is nonlinear; if the frequencies of two sound waves differ by a factor of 2, their pitches are said to be an *octave* apart. Sounds at 256 Hz and 512 Hz are one octave apart, but so are sounds at 3 kHz and 6 kHz. The perceived pitch also depends on the amplitude of the sound wave. Above 3 kHz, pitch increases with increasing sound wave amplitude; below 2 kHz, pitch decreases with increasing amplitude.

Timbre: We can readily distinguish two musical instruments even when they emit tones of identical pitch because they produce different mixtures of harmonics. The pitch is typically related to the lowest, or fundamental, frequency in the series of harmonics, and two instruments playing the same pitch usually produce the same fundamental. They differ, however, in the content of the higher harmonics, which is what determines the timbre of the instrument. Even though the ear and the brain constitute a remarkable Fourier analyzer in the sense that they can distinguish sounds of different harmonic contents, we cannot characterize timbre with a precise statement about measurable parameters. (An alternative term you might run across for timbre is *tonal quality*.)

17.4 (*a*) Plot the velocity of the beads along the chain in Figure 17.9*b* as a function of their equilibrium position x. (*b*) Plot the linear density (number of beads per unit length) as a function of x.

As Checkpoint 17.4 suggests, we can also represent the wave in Figure 17.9 by plotting the linear density versus position. Comparing the linear density curve of Figure S17.4*b* with the $D_x(x)$ curve of Figure S17.4*a* (see the Solutions to Checkpoints near the back of this book) gives us some important information:

The compressions and rarefactions in longitudinal waves occur at the locations where the medium displacement is zero.

(You can get this same information by comparing Figures 17.9*b* and *c*.)

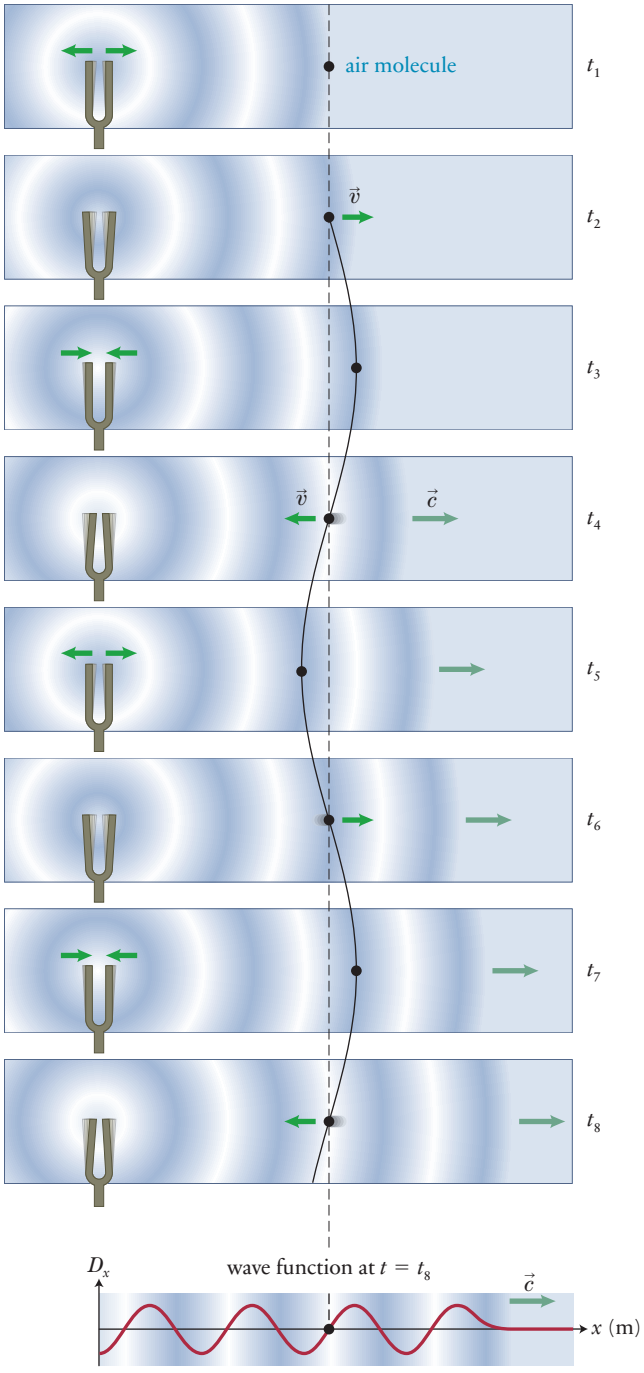
The propagation of sound through a medium is very similar to that of a compressional wave through the chain in Figure 17.9. Disturbed molecules in a solid, liquid, or gas are subject to intermolecular forces that tend to return the molecules to their equilibrium positions. For small displacements, the restoring force is proportional to the displacement. It is this restoring force that is responsible for the propagation of the disturbance caused by the piston in Figure 17.8.

When not confined to a tube, a compressional wave traveling through air spreads out in three dimensions because, unlike beads along a string, air molecules interact with the surrounding molecules in all three directions. Consequently, sound waves tend to form spherical wavefronts. **Figure 17.10** shows a sound wave generated by an oscillating tuning fork. A series of spherical shells of compression and rarefaction travel outward from the fork. Each molecule in the air executes a sinusoidal back-and-forth movement as successive shells travels by; the sine curve in the figure shows the position versus time of one molecule.

The displacements of air molecules disturbed by a sound wave are very small. At a distance of a few hundred meters, for example, the deafening roar of an airliner taking off causes displacements of only 10^{-6} m—about one-hundredth of the width of a hair. These small displacements coupled with the high speed at which sound waves travel make it impossible to directly see the compressional wavefronts that accompany sound. Only under special laser illumination is it possible to observe these wavefronts (**Figure 17.11**).

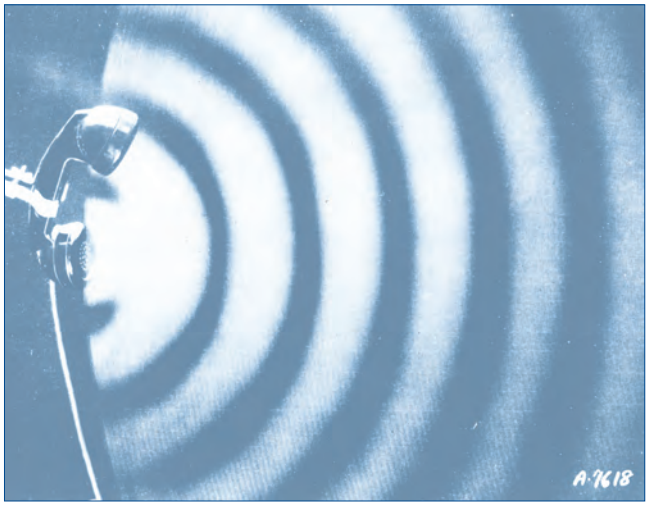
Notice how the wavefronts shown in Figures 17.10 and 17.11 resemble the circular wavefronts of a surface wave, such as the one represented in Figure 17.2. The compressions correspond to crests, and the rarefactions correspond to troughs. When we draw circles to represent the spherical wavefronts of sound waves, which are longitudinal waves, the circles represent not crests and troughs but rather compressions and rarefactions.

Figure 17.10 The periodic wave emitted by a tuning fork causes a series of outward-traveling spherical wavefronts. The dot represents an air molecule that oscillates about its equilibrium position as the wavefronts pass by.



CONCEPTS

Figure 17.11 Image of the spherical compressional wavefronts produced by sound waves emanating from a telephone. This image was obtained using laser illumination. *Reused with permission of Nokia Corporation and AT&T Archives.*



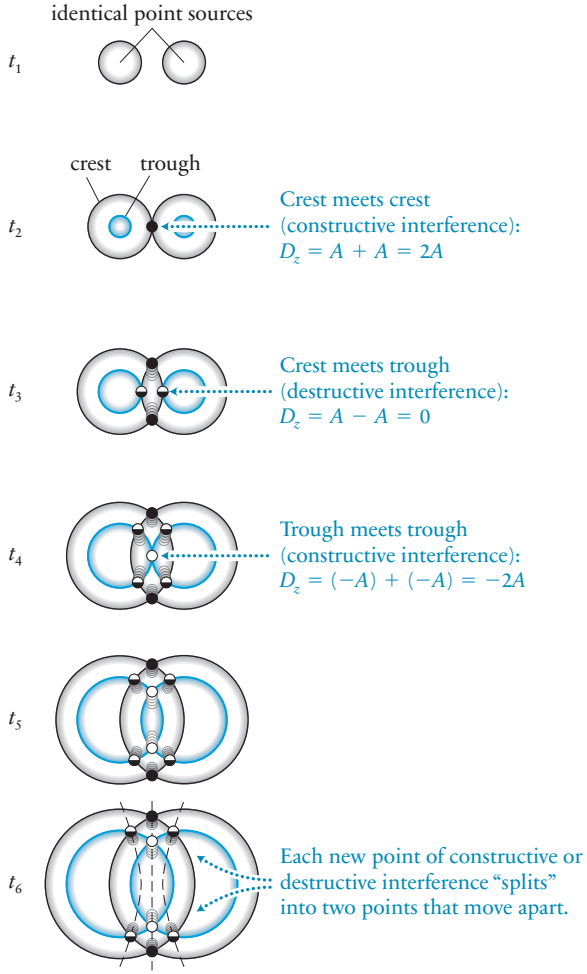
17.5 Even though the sound of a common type of loudspeaker carries in all directions, the loudness in front of the speaker is considerably greater than the loudness behind it. What is the shape of the wavefronts produced by such a loudspeaker?

17.3 Interference

The superposition of overlapping waves in two and three dimensions leads to a number of interesting phenomena. Consider first the situation illustrated in **Figure 17.12**, where two identical circular wave pulses traveling on a liquid surface cross as they expand. At t_2 the two crests meet at a point midway between the two sources, giving rise to constructive interference. At this point the displacement of the liquid surface is twice as great as that due to the individual pulses. As the pulses continue to expand, this point of overlap splits into two points, one moving upward and the other moving downward. At t_3 the crest of each pulse meets the trough of the other pulse, leading to destructive interference. Because the height of the crest is equal to the depth of the trough, the vector sum of the medium displacement is zero.* Each of these points splits into two points that move vertically away from each other. At t_4 the two troughs meet and reinforce each other, giving rise to constructive interference again. At this point the depression of the medium is

*The waves do not cancel exactly close to the sources because the amplitudes of the two waves differ due to the $1/\sqrt{r}$ factor in the wave amplitude. Far from the sources, however, this difference is negligible.

Figure 17.12 Interference of two identical sets of circular wave pulses expanding on a liquid surface.

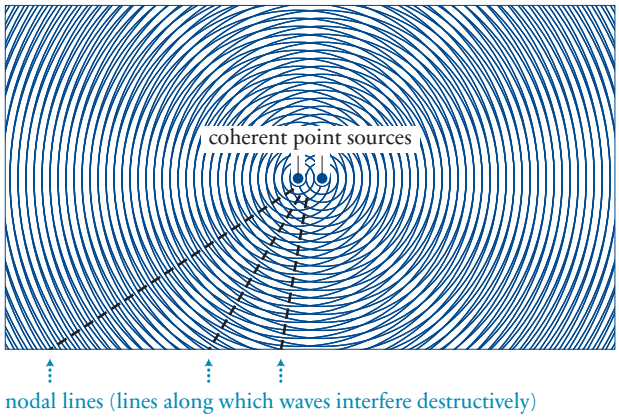


doubled. Note that, as shown at t_6, the points of constructive interference move along a straight line perpendicular to the line that connects the two sources, while the points of destructive interference move along curved paths.

Let us now apply these ideas to the overlapping of periodic surface waves. When periodic circular waves from two nearby sources cross each other, the crests and troughs overlap in many places, giving rise to a complex interference pattern. **Figure 17.13** on the next page shows the circular wavefronts for two sources separated by a distance equal to three wavelengths and emitting waves of the same amplitude and frequency. In addition, the two sources always have the same phase and are therefore said to be *in phase*. Sources that are emitting waves that have a constant phase difference are said to be **coherent.** The pattern produced by the overlapping circles in Figure 17.13, called a *moiré pattern*, creates the illusion of a number of bands radiating out from the center of the figure (*moiré* is the French word for the wavy finish seen in some fabrics). As we shall see shortly,

CONCEPTS

Figure 17.13 Moiré pattern formed by the overlapping wavefronts of two closely spaced, coherent wave sources (located at the centers of the two concentric sets of circles). The center of each band in the pattern corresponds to a nodal line.



the centers of these bands correspond to **nodal lines**—lines where the two waves cancel each other and the vector sum of the medium displacement is always zero.

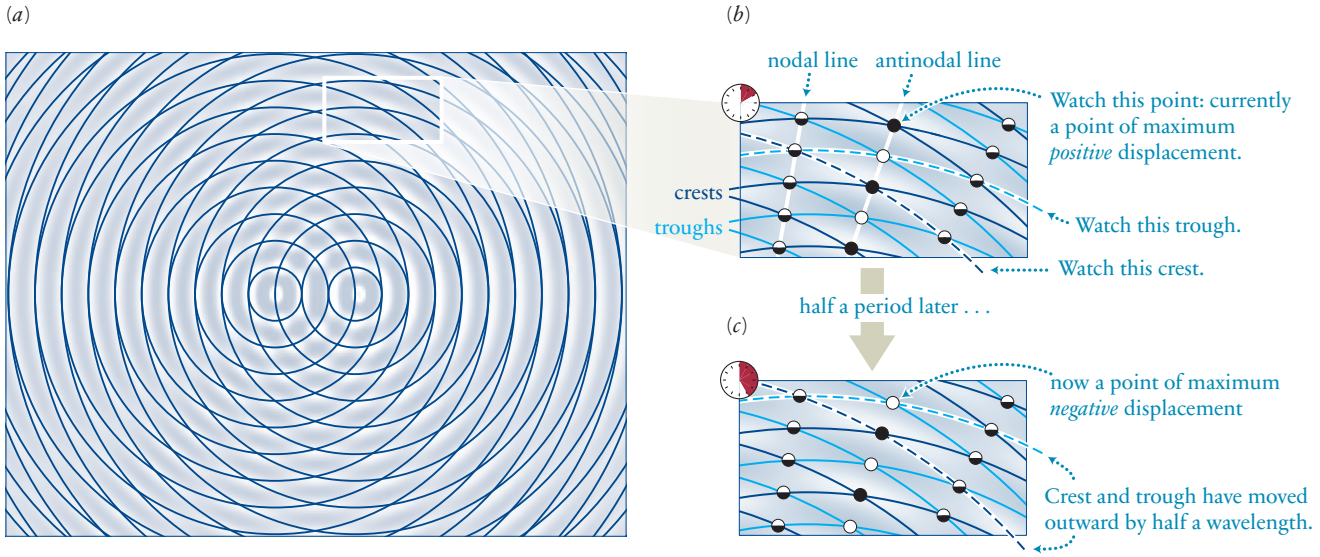
Figure 17.14*a* shows a magnified view of the central portion of Figure 17.13 overlaid on a picture of the interference pattern. White areas in the picture correspond to areas where the medium displacement is greatest—regions where two crests reinforce each other (comparable to t_2 in Figure 17.12). Bright blue areas also correspond to areas where the medium displacement is greatest—regions where two troughs reinforce each other (t_4 in Figure 17.12). Pale blue areas correspond to regions where the surface is undisturbed because the waves interfere destructively. If you look at the picture from a glancing angle, you can see that these pale blue areas form outward radiating bands. The bands coincide with the positions where the wavefronts of one source fall halfway between the wavefronts of the other (which gives the illusion of bands in the moiré pattern in Figure 17.13).

Figure 17.14*b* shows a detail of the interference pattern. In addition to the solid dark blue arcs that indicate wavefronts marking the crests of the waves generated by each source, bright blue arcs indicate the locations of the troughs halfway between each pair of neighboring crests. Figure 17.14*c* shows the same region half a period later. All wavefronts have advanced by half a wavelength, and the nodes (◒) and antinodes (● and ○) have moved along with the wavefronts. See, for example, the node at the intersection of the dashed dark blue crest from the left source and the dashed bright blue trough from the right source—in Figure 17.14*b* it is the second node from the top; in Figure 17.14*c* it is the uppermost. In addition, all white and bright blue areas are interchanged, as are all filled and unfilled circles (see, for example, the highlighted filled and open circles in Figure 17.14*b* and *c*). Where two crests initially reinforced each other, two troughs now meet. (Of course, the nodal lines stay in place, by definition.)

17.6 (*a*) What is the medium displacement at a point halfway between a filled circle and a neighboring open circle in Figure 17.14*b*? (*b*) Sketch the variation of this displacement as a function of time. (*c*) What is the medium displacement at all points between two neighboring half-filled circles? (*d*) Sketch the variation of this displacement as a function of time at the midpoint between two neighboring half-filled circles.

CONCEPTS

Figure 17.14 Details of the interference pattern in the central region of Figure 17.13. White and blue regions represent, respectively, areas of maximum positive and maximum negative medium displacement. The intermediate regions of zero displacement line up to form the gray nodal bands seen in Figure 17.13.



The half-filled circles line up to form a path, and as Checkpoint 17.6 demonstrates,* *all* points along this path remain undisturbed and are therefore nodes. The line that connects the half-filled circles is thus a nodal line, and the pale blue bands that appear in Figure 17.13 are thus centered on nodal lines. The filled and open circles line up to form what are called **antinodal lines.** Points on these antinodal lines have the greatest amplitude of oscillation; in other words, these points are antinodes.

17.7 Is a point halfway between a filled and a neighboring open circle in Figure 17.14*b* a node?

The interference pattern formed by overlapping waves in two dimensions can readily be observed for waves generated on a water surface (**Figure 17.15**). The nodal lines are clearly visible in the picture. All waves in more than one dimension—sound waves, light waves, x-ray waves, radio waves—set up interference patterns that contain nodal and antinodal lines. For example, two loudspeakers that are some distance apart and that emit sound waves of the same frequency set up an interference pattern that is not very different from the ones for circular water waves. By adjusting your position in front of the speakers, you can find regions where the destructive interference occurs, causing the loudness of the sound to be low.

The existence of nodal regions has a number of consequences. Consider, for example, the interference pattern shown in **Figure 17.16*a***. Two coherent sources S_1 and S_2 emit circular waves. The waves cause the medium displacement at point P to vary, but at point Q, located on a nodal line, the medium displacement *remains zero*. If you were standing at that location, you could not tell whether the sources were emitting waves or not! Each source separately, however, does emit waves in the direction of Q. Figure 17.16*b* shows that, with S_1 is turned off, the wavefronts from S_2 that reach Q are not interfered with. In other words, if you were standing at Q, you would hear sound when S_1 is turned off but not when both sources are on. So, *removing* a source *increases* the amplitude of the wave at Q! Conversely, adding a second source (in other words, returning to the situation in Figure 17.16*a*) decreases (or even eliminates) the wave amplitude in certain regions. So:

Figure 17.15 Photograph of the interference pattern caused by overlapping circular waves on a water surface.

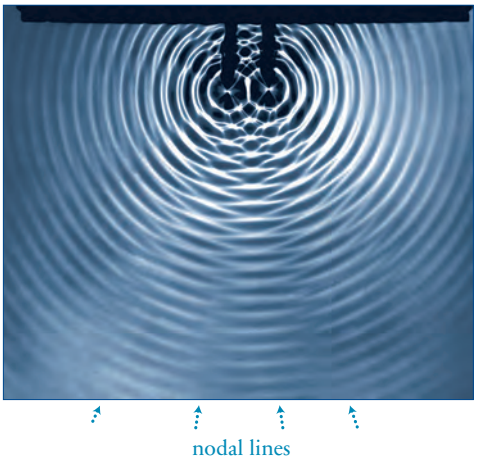
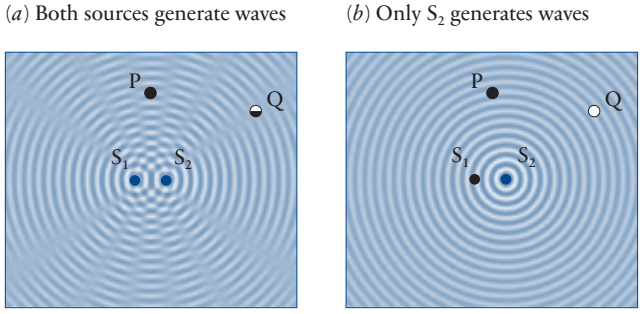


Figure 17.16 Less can be more. (*a*) When both sources generate waves, point Q lies on a nodal line and thus never experiences displacement of the medium (because the waves from the two sources interfere destructively there). If these were sound waves, there would be silence at Q. (*b*) However, when only one source generates waves, Q experiences their full effect.



When the waves from two coherent sources interfere, the amplitude of the sum of these waves in certain directions is less than that of a single source.

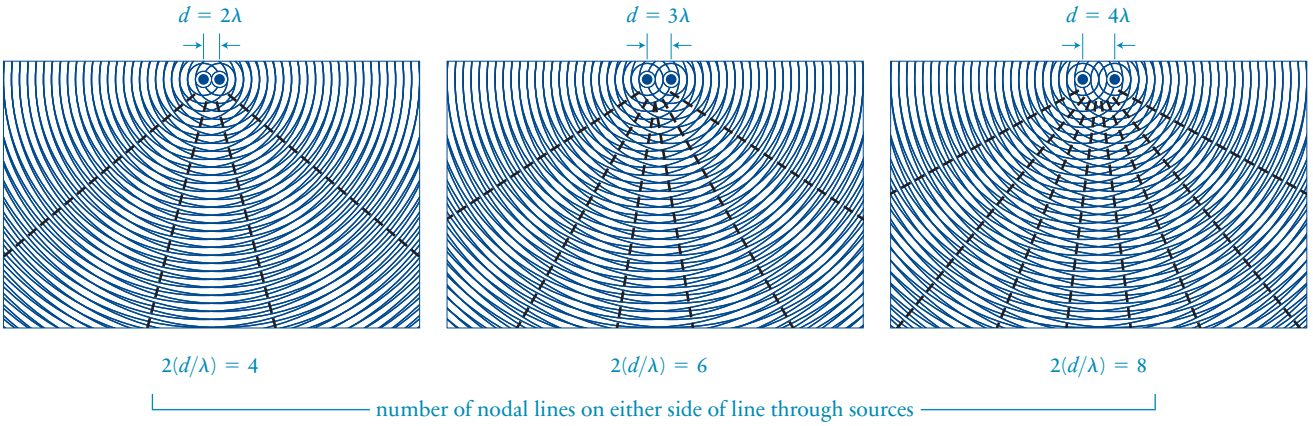
17.8 (*a*) If the amplitude of the wave generated by each source in Figure 17.16 is *A* at points P and Q, what is the overall amplitude at P in Figure 17.16*a*? In Figure 17.16*b*? (*b*) Given that the wave energy that passes Q in Figure 17.16*b* becomes zero when S_1 is turned on, do you expect the energy that passes P to increase or decrease? (*c*) Use your answer to part *a* to determine by what factor the amount of energy that passes P changes, and reconcile your answer with your answer to part *b*.

Just what a given interference pattern looks like is determined by the wavelength and the spacing between the two sources. The greater the source separation relative to the wavelength, the greater the number of nodal lines. We determine the number of nodal lines in a pattern by imagining a straight line running through

CONCEPTS

*I cannot overstress the importance of doing Checkpoint 17.6. By completing this checkpoint, you will be teaching yourself the key feature of interference of waves in more than one dimension.

Figure 17.17 The effect of changing the separation between two point sources.



the centers of the two sources and then counting the number of nodal lines on either side of this line:

> **If two coherent sources located a distance d apart emit identical waves of wavelength λ, then the number of nodal lines on either side of a straight line running through the centers of the sources is the greatest integer smaller than or equal to $2(d/\lambda)$.**

You can study the appearance of nodal lines as the separation between the sources is varied in **Figure 17.17**.

17.9 By how many wavelengths are the sources in Figure 17.16*a* separated?

EXAMPLE 17.3 Nodes

For the situation shown in Figure 17.16*a*, how many nodes are there along the line segment S_1S_2 that connects the centers of the two sources?

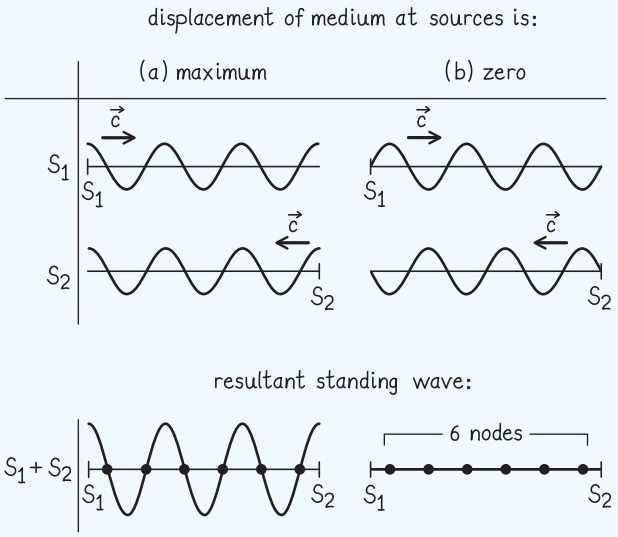
1 GETTING STARTED The sources are coherent and so send identical waves in opposite directions along S_1S_2. This means there is a standing wave (see Section 16.6) along that segment. From Figure 17.16*b* I can tell that exactly three wavelengths fit between S_1 and S_2.

2 DEVISE PLAN The two sources are in phase, so I can determine the number of nodes by plotting the amplitudes of the waves due to each source along S_1S_2 at two instants and then adding those amplitudes. The nodes occur wherever the amplitude of the combined wave remains zero.

3 EXECUTE PLAN **Figure 17.18*a*** shows, for the instant at which the medium displacement at both sources is maximum, the medium displacement caused by S_1 and S_2 and the sum of the two. Figure 17.18*b* shows the displacements one-quarter cycle later. Comparing the two graphs, I see that six positions along line segment S_1S_2 (denoted by the black dots in the bottom graphs) remain stationary, indicating six nodes. ✔

4 EVALUATE RESULT Because the sources are separated by a distance equal to three wavelengths and because there is one node for each half-wavelength (see Section 16.6), my answer makes sense.

Figure 17.18



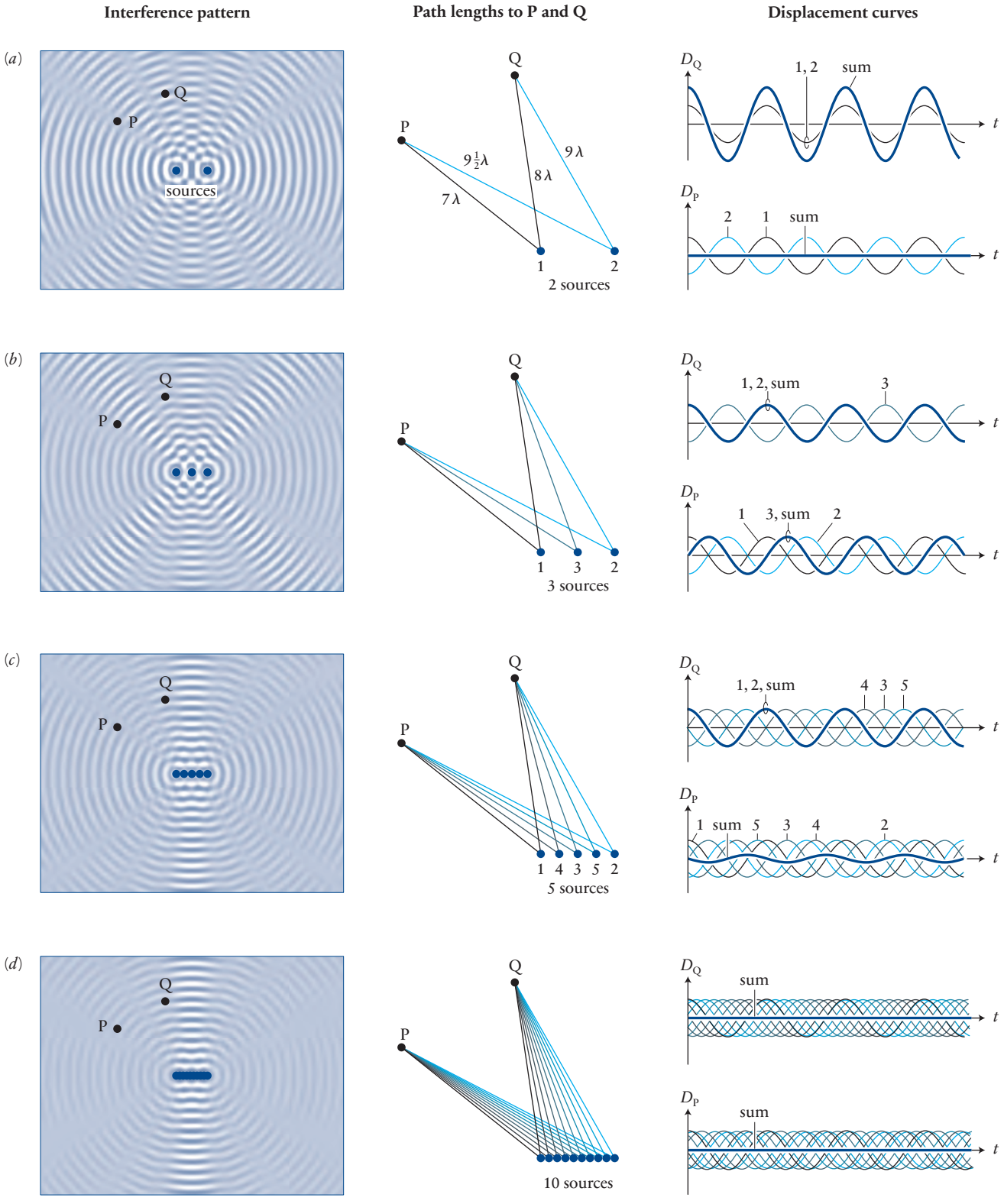
Example 17.3 shows us why there are twice as many nodes and therefore twice as many nodal lines as there are wavelengths between the sources. Each node corresponds to a nodal line, and so you now see why there is a connection between the number of nodal lines and the source spacing.

17.10 Along the leftward and rightward extensions of the straight line joining S_1 and S_2 in Figure 17.16, why are there no nodes to the left of S_1 and to the right of S_2?

Figure 17.19 shows what happens when additional coherent sources are added between the two sources of Figure 17.14. Let us focus our attention on the two points P and Q. When just two sources are present (Figure 17.19*a*), P is on a nodal line and Q is on an antinodal line. The path length diagram in Figure 17.19*a* shows the waves arriving at P and Q.

CONCEPTS

Figure 17.19 When additional coherent sources are added between the two sources spaced by three wavelengths, the interference pattern changes: The greater the number of sources, the more the waves are concentrated along the axis perpendicular to the sources. The displacement curves show the displacements at P and Q for each source and the sum of these displacements (bold curve).



The distance from source 1 to P is 7λ; the distance from source 2 to P is $9\frac{1}{2}\lambda$. Consequently the two waves *always* arrive at P out of phase—they cause medium displacements in opposite directions and so always cancel. The distance from source 1 to Q is 8λ; the distance from source 2 is 9λ. So the two waves arrive in phase at Q and always reinforce each other.

Figure 17.19*b* shows how the addition of a single source halfway between the two original sources redistributes the flow of wave energy and concentrates the waves in fewer triangular areas. The phase of the wave arriving from the additional source is approximately halfway between that from sources 1 and 2. In other words, source 3 is about $8\frac{1}{4}\lambda$ from P and $8\frac{1}{2}\lambda$ from Q. Because the waves from 1 and 2 cancel each other at P, the wave from 3 arrives undisturbed at P, so P is no longer on a nodal line. At Q, the additional wave from 3 arrives out of phase with those from 1 and 2, so the additional source diminishes the wave amplitude at Q.

If two additional sources are added, almost all the wave energy becomes concentrated in a narrow cone perpendicular to the straight line that connects the sources (Figure 17.19*c*). The phases of the waves are such that, even though there are five sources instead of two, the amplitude of the waves arriving at Q is *smaller* than before and that of the waves arriving at P is very small.

With ten coherent sources, all the waves are concentrated in a cone along the direction perpendicular to the straight line that joins the sources (Figure 17.19*d*). Outside of this cone, the waves—all arriving at any given point, including P and Q, with a different phase—cancel each other. **Figure 17.20** shows what happens when 100 coherent sources are placed close together, now forming a row 10λ wide: A *beam* of nearly straight wavefronts emanates in a direction perpendicular to the row of sources. There are hardly any waves outside the beam because the waves from all the different sources tend to cancel each other. Summarizing:

When many coherent point sources are placed close together along a straight line, the waves nearly cancel out in all directions except the direction perpendicular to the axis of the sources.

17.11 How does the wave amplitude along the beam of wavefronts in Figure 17.20 change with distance from the row of sources?

17.4 Diffraction

The nearly straight wavefronts that emanate from continuous rows of coherent sources suggest that *any* wavefront may be regarded as having originated from a collection of many closely spaced, coherent point sources emitting waves in all directions. The waves cancel in all directions except the direction perpendicular to the wavefront. This picture of wavefronts consisting of closely spaced, coherent point sources is called **Huygens' principle** after the 17th-century Dutch physicist Christiaan Huygens.

Figure 17.20 A 10λ-wide row of 100 coherent sources produces nearly planar wavefronts. The flow of energy is entirely along a beam perpendicular to the row of sources.

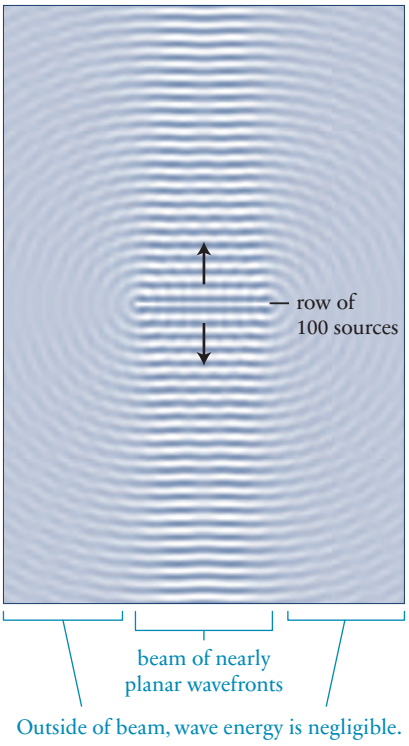


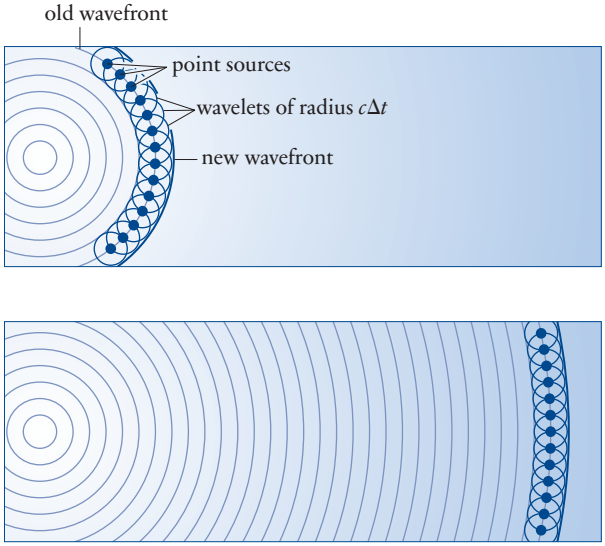
Figure 17.21 illustrates how Huygens' principle can be used to construct a future wavefront given the wavefront at a certain instant. We begin by dividing the original wavefront into a series of closely spaced point sources. All those point sources emit in phase because they are all on the same wavefront. After a short time interval Δt, each source has emitted a small circular wavefront, called a *wavelet*, of radius $c\Delta t$. Wavelets of neighboring sources overlap in such a way that they interfere destructively with each other except in the forward direction. All these forward-moving wavelets combine to form the wavefront we aim to construct, labeled "new wavefront" in Figure 17.21.*

Huygens' principle also helps explain how waves spread around corners. If you stand near an open window, you hear sounds from the street even if you are

*For a *single* row of coherent point sources, wavefronts emanate in *two* directions along a line perpendicular to the straight line connecting the sources (Figure 17.20). When additional rows of coherent point sources are added in front of or behind this row and these sources are given the proper phases, the amplitude of the resulting waves in one of the two directions is suppressed by interference.

CONCEPTS

Figure 17.21 According to Huygens' principle, any wavefront can be regarded as a collection of closely spaced, coherent point sources. Advancing wavefronts can then be constructed by drawing circular wavelets centered at points along an earlier wavefront.



not standing directly in front of the window. Likewise, if you stand behind a tree, you can hear sounds coming from in front of the tree because the sound waves spread around the tree.

Figure 17.22 shows nearly planar wavefronts incident on gaps of various size. If the gap is much larger than the wavelength, the gap lets through a *beam* of wavefronts with very little spreading. As the gap gets narrower, the waves beyond the gap spread out more and more, so that what gets through the gap no longer constitutes a beam. When the gap width is smaller than the wavelength, the wavefronts beyond the gap spread out in all directions—in essence, the gap now acts like a single point source. This spreading, called **diffraction,** is another characteristic feature that distinguishes waves from objects: Waves spread out when going through a narrow opening, but objects do not. If you throw a handful of pebbles through an opening, for instance, the "beam" of pebbles doesn't spread out after passing the opening.

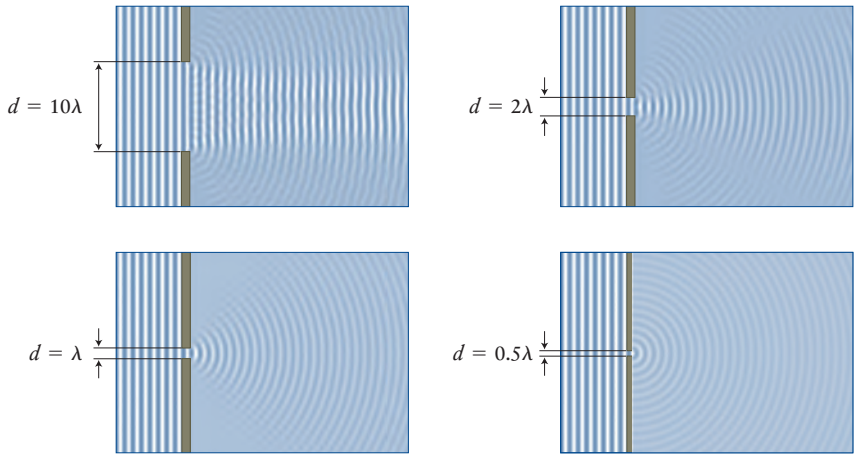
The reason waves diffract is directly related to the observation that many coherent point sources in a row give rise to waves in just one direction because the waves emitted by each source cancel in all but the forward direction (Figure 17.20). Here we see the opposite occurring: As the number of point sources is reduced (the gap is narrowed in our example), the cancellation of waves along directions away from the straight line perpendicular to the line of sources becomes less complete and spreading occurs. As the width of the gap becomes comparable to the wavelength, the gap acts like a point source, leading to spreading of the wave in all directions.

Obstacles or apertures (openings) whose width is smaller than the wavelength of an incident wave give rise to considerable spreading of that wave.

The amount of spreading depends strongly on the wavelength. For a given aperture, low-frequency sounds diffract more strongly than higher-frequency ones. If a band marches by below your window and you are not standing directly in the window opening, you hear the low-frequency drone of drums and tubas more than the high-frequency sounds of cymbals and piccolos. Move near the window and the sound of the band turns from dull to sharp as the high-frequency sounds emitted by it now also reach your ears.

CONCEPTS

Figure 17.22 When a planar wavefront passes through a gap, some spreading, called diffraction, occurs. The smaller the gap, the greater the amount of spreading. When the width d of the gap is smaller than the wavelength λ, the gap acts like a point source.



A direct consequence of wave diffraction is that it is impossible to create a beam of waves that is narrower than the wavelength of the waves. This means that diffraction limits the ability of waves to locate objects. For example, it is very difficult for a person to locate small objects by detecting reflected sound waves because, with their relatively long wavelengths (0.02 m to 20 m), these waves diffract around most objects. We therefore locate objects using visible light, whose wavelength is in the range of 4×10^{-7} m to 7×10^{-7} m (less than one-hundredth the width of a human hair). Only objects and apertures that are smaller than the wavelength of visible light diffract light.

Bats can detect ultrasonic waves of frequencies up to 120 kHz. The wavelength of these waves is about 3 mm. By emitting such waves and detecting their reflections, bats can locate small insects. Ultrasonic waves are also used for diagnostic purposes in medicine. Because of their short wavelength, it is possible to emit a narrow beam of ultrasonic waves. The beam is rapidly scanned across a sample to create an image from the reflections of the beam that occur inside the sample (**Figure 17.23**).

17.12 Suppose the barriers in Figure 17.22 were held at an angle to the incident wavefronts. Sketch the transmitted wavefronts for the case where the width of the gap is much smaller than the wavelength of the incident waves.

Figure 17.23 Ultrasound image of a three-month-old fetus in utero. The 2-MHz waves (1 megahertz = 1 MHz = 10^6 Hz) used to obtain this picture have a wavelength of 0.7 mm.

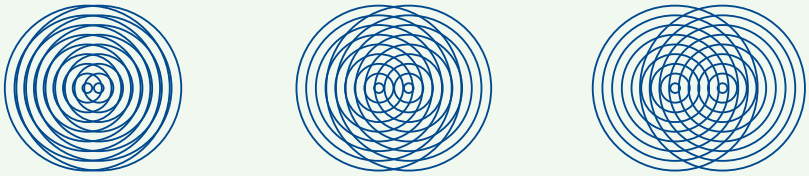


CONCEPTS

Self-quiz

1. A circular wave travels such that a crest occurs at radius r_1 at instant t_1. The same crest has a radius $r_2 = 2r_1$ at $t_2 = 2t_1$. (*a*) What is the wave speed? (*b*) Where is the crest at $t_3 = 3t_1$? (*c*) What is the ratio (energy per unit length at t_3)/(energy per unit length at t_1)?
2. A spherical wave travels such that a crest occurs at radius r_1 at instant t_1. The same crest has a radius $r_2 = 2r_1$ at $t_2 = 2t_1$. (*a*) What is the wave speed? (*b*) Where is the crest at $t_3 = 3t_1$? (*c*) What is the ratio (energy per unit area at t_3)/(energy per unit area at t_1)?
3. Two circular wavefronts interfere as shown in **Figure 17.24**. (*a*) Draw the antinodal lines. (*b*) Is the medium displacement along these lines constant?

Figure 17.24

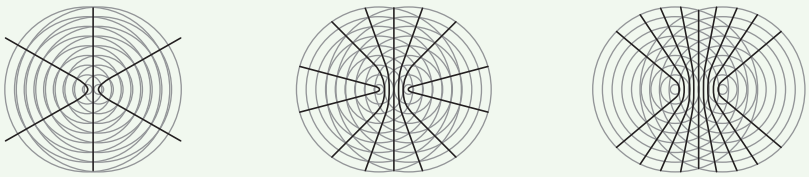


4. Two coherent sources are placed two wavelengths apart. Point P lies on a nodal line in the interference pattern of the two sources, and point Q lies on an antinodal line. (*a*) If one source is turned off, does the energy arriving at P increase, decrease, or remain the same? (*b*) If one source is turned off, does the energy arriving at Q increase, decrease, or remain the same? (*c*) Do your answers depend on which source is turned off?
5. Because sound waves diffract around an open doorway, you can hear sounds coming from outside the doorway. You cannot, however, see objects outside the doorway unless you are directly in line with them. What does this observation imply about the wavelength of light?

Answers

1. (*a*) The speed is the distance traveled divided by the time interval over which the motion occurs: r_1/t_1. (*b*) At $3t_1$ the crest is at $3r_1$. (*c*) The energy per unit length at r_1 is proportional to $1/r_1$. That at r_3 is proportional to $1/r_3 = 1/3r_1$. The ratio of the energies per unit length at the two instants is thus $1/3$.
2. (*a*) The speed is r_1/t_1. (*b*) The crest is at $3r_1$. The answers to parts *a* and *b* to this and the preceding question have to do with only the speed of the waves, not their dimensionality. (*c*) The energy per unit area at r_1 is proportional to $1/r_1^2$. That at r_3 is proportional to $1/r_3^2 = 1/(3r_1)^2$. The ratio of the energies per unit area at the two instants is $1/9$.
3. (*a*) See **Figure 17.25**. (*b*) No. The medium displacement at any point on an antinodal line varies as a function of time.

Figure 17.25



4. (*a*) Being on a nodal line, P initially receives no energy. After one source is turned off, it receives energy from the other source, and so the energy P receives increases. (*b*) Being on an antinodal line, Q receives four times the energy of a single source because the amplitude at Q is twice that of a single source. When one source is turned off, Q receives only the energy from a single source, and so the energy received at Q decreases. (*c*) It does not matter which source is turned off. The interference pattern exists only when both sources are on, and each point receives the energy from only a single source when the second source is turned off.
5. Because light does not diffract as it travels through the doorway, this observation implies that the wavelength of the light must be smaller than the width of the doorway. Given that visible light has wavelengths between 4×10^{-7} m and 7×10^{-7} m and most doorways are about 1 m wide and 2 m tall, this is indeed the case.

17.5 Intensity

As we saw in Section 17.1, the propagation of waves in more than one dimension causes the wave energy to be diluted over a larger and larger region. For waves in three dimensions, we define the **intensity** I as the energy delivered by the wave per unit time per unit area normal to the direction of propagation. Because the energy delivered per unit time is the definition of power, we can say

$$I \equiv \frac{P}{A}, \tag{17.1}$$

where P is the power delivered by the wave over an area A. For three-dimensional waves, the SI unit of intensity is thus $\mathrm{W/m^2}$.

Consider a point source sending out waves uniformly in all directions. If there is no dissipation of energy, all the energy generated by the source must cross an imaginary sphere centered on the source. If the power delivered by the source is P_s, the intensity is

$$I = \frac{P_s}{A_{\text{sphere}}} = \frac{P_s}{4\pi r^2} \quad \text{(uniformly radiating point source)}, \tag{17.2}$$

and so we see that the intensity of a wave in three dimensions is proportional to $1/r^2$, as we concluded in Section 17.1.

For a two-dimensional surface wave, the intensity is defined as the energy incident per unit time per unit *length* normal to the direction of propagation. For a surface wave that delivers a power P over a length L,

$$I_{\text{surf}} \equiv \frac{P}{L}. \tag{17.3}$$

Consequently the SI unit of the intensity of a surface wave is W/m because we must replace the area of a sphere in Eq. 17.2 by the circumference of a circle: $I = P_s/(2\pi r)$. The intensity of two-dimensional waves is thus proportional to $1/r$.

EXAMPLE 17.4 Sound wave energy

The minimum intensity audible to the human ear is called the *threshold of hearing*. For a 1.0-kHz sound wave, this threshold is approximately 10^{-12} $\mathrm{W/m^2}$. The maximum tolerable intensity, called the *threshold of pain*, is about 1.0 $\mathrm{W/m^2}$ for a 1.0-kHz sound wave. (*a*) For a 1.0-kHz sound wave, estimate the amount of energy delivered to the ear in 1.0 s at the threshold of hearing and at the threshold of pain. (*b*) If the sound is produced by a loudspeaker 1.0 m away from the ear, what is, in each case, the power emitted by the loudspeaker?

❶ **GETTING STARTED** I am given two intensities, which give the power per unit area, and asked in part *a* to calculate the energy arriving at the ear in 1.0 s for each intensity. Because energy per unit time is power, part *a* reduces to identifying the area over which the energy is delivered. For part *b*, I am given a loudspeaker-to-ear distance, and I need to calculate the power for each loudspeaker at that distance.

❷ **DEVISE PLAN** For part *a*, to estimate the energy delivered to the ear in 1.0 s, I have to estimate the surface area of the ear. I can then use Eq. 17.1 to calculate the power delivered to the ear. To obtain the energy delivered, I multiply the power by the time interval. For part *b*, I assume that the loudspeaker is a point source that radiates sound uniformly in all directions. With this assumption, I can use the given intensities and source-to-ear distance in Eq. 17.2 to calculate the power emitted by the speaker.

QUANTITATIVE TOOLS

3 EXECUTE PLAN (*a*) The surface area of the outer ear is approximately $(0.030\ \text{m})^2 \approx 9.0 \times 10^{-4}\ \text{m}^2$, so the power delivered to the ear at the threshold of hearing is about

$$P = IA = (10^{-12}\ \text{W/m}^2)(9.0 \times 10^{-4}\ \text{m}^2) = 9.0 \times 10^{-16}\ \text{W}$$

$$= 9.0 \times 10^{-16}\ \text{J/s}.$$

In 1.0 s, a mere 9.0×10^{-16} J—about a millionth of a billionth of a joule—gets delivered to the ear when a sound source emits at an intensity equal to the threshold of hearing. ✔

At the threshold of pain, the power delivered to the ear is

$$(1.0\ \text{W/m}^2)(9.0 \times 10^{-4}\ \text{m}^2) = 9.0 \times 10^{-4}\ \text{W},$$

and so 0.90 mJ is delivered to the ear in 1.0 s. ✔

(*b*) For the power emitted by the loudspeaker at the two thresholds, Eq. 17.2 gives

threshold of hearing:

$$P_s = I(4\pi r^2) = (10^{-12}\ \text{W/m}^2)[4\pi(1.0\ \text{m})^2]$$

$$= 1.3 \times 10^{-11}\ \text{W}\ ✔$$

threshold of pain:

$$P_s = (1.0\ \text{W/m}^2)[4\pi(1.0\ \text{m})^2] = 13\ \text{W}.\ ✔$$

4 EVALUATE RESULT The energies delivered at the threshold of hearing are very small, and I have no good way of evaluating those. I do know that the ear can detect sounds made by the tiniest of movements—the rustle of a leaf, for example—and these movements cannot transmit much energy, so at least I know that the values I obtained are not unreasonably large. I know that headphones and loudspeakers can easily emit power levels at the threshold of pain. The fact that headphones can be powered with tiny batteries tells me that the 0.9-mW answer I obtained in part *a* is not unreasonable. Most stereo systems are rated at tens to hundreds of watts, so the 13 W in part *b* is also in the right range.

As Example 17.4 illustrates, the human ear is capable of handling sound waves over a tremendous range of intensities—intensities that differ by as much as a factor of 10^{12}. To deal conveniently with such an enormous intensity range, it is common to use a logarithmic scale for intensity. For example, if the intensity I increases by a factor of 10^{12}, as it does when going from the threshold of hearing to the threshold of pain, the common logarithm (see Appendix B), $\log I$, increases by only 12. The intensity of a sound wave can then be expressed as part of a ratio that also includes the intensity I_{th} at the threshold of hearing:

$$\log\left(\frac{I}{I_{\text{th}}}\right). \tag{17.4}$$

The reference level ($I_{\text{th}} = 10^{-12}\ \text{W/m}^2$) is chosen to coincide with human hearing, and the quantity $\log(I/I_{\text{th}})$, even though it is unitless, is given an artificial unit. This unit is the *bel* (B), in honor of the American scientist Alexander Bell (1847–1922). It is customary to express intensity levels of sound in **decibels** (dB), one-tenth of a bel: 1 dB = 0.1 B. The **intensity level** β of a sound, measured in decibels, is thus defined as

$$\beta \equiv (10\ \text{dB}) \log\left(\frac{I}{I_{\text{th}}}\right). \tag{17.5}$$

Figure 17.26 Average auditory response of the human ear. The ear is sensitive to longitudinal waves in air ranging in frequency from 20 Hz to 20 kHz. The sensitivity is greatest at 3 kHz, where the threshold of hearing is lowest. The curves give the intensity levels required to produce harmonic tones of the same perceived loudness.

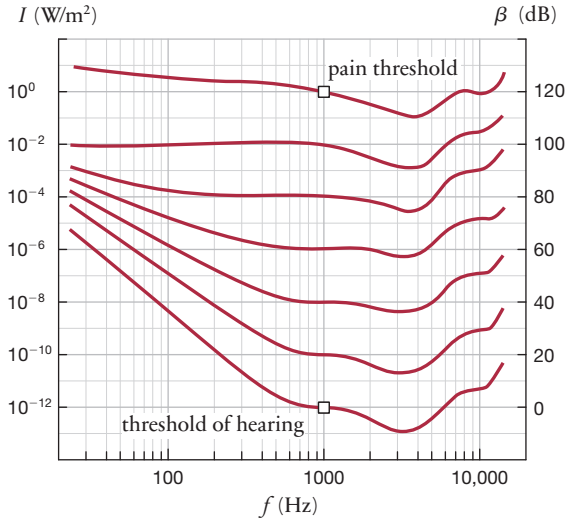


Table 17.1 Approximate intensity levels

Source	distance (m)	β (dB)	Description
Jet engine	50	140	pain
Pneumatic hammer	10	110	
Shout	1.5	100	very loud
Car horn	10	90	
Hair dryer	0.2	80	loud
Automobile interior		70	
Conversation	1	60	moderate
Office background		50	
Library background		40	
Suburban bedroom		30	quiet
Whisper	1	20	
Normal breathing	5	10	barely audible

At the threshold of hearing, where the intensity is $I = 10^{-12}\ \text{W/m}^2$, we have $\log(I/I_{\text{th}}) = \log(1) = 0$, and so the intensity level $\beta = (10\ \text{dB}) \times (0) = 0\ \text{dB}$. At the threshold of pain, where the intensity is $I = 1\ \text{W/m}^2$, we have $\log[(1\ \text{W/m}^2)/(10^{-12}\ \text{W/m}^2)] = 12$, and so the intensity level is $\beta = (10\ \text{dB})(12) = 120\ \text{dB}$.

As Figure 17.26 illustrates, the thresholds for hearing and pain vary with frequency. The five curves between the two threshold curves give the intensity required to produce harmonic sounds that a listener perceives as all being equally loud. For example, a sinusoidal 200-Hz tone of intensity $10^{-8}\ \text{W/m}^2$ and a sinusoidal 1-kHz tone of intensity $10^{-10}\ \text{W/m}^2$ are perceived to have the same loudness by the average listener. The right vertical scale gives the corresponding intensity level in decibels. Table 17.1 gives the intensity levels in decibels of a number of common sound sources.

EXERCISE 17.5 Doubling the intensity

A clarinet can produce sound waves of intensity level about 70 dB. By how much does the intensity level increase if a second clarinet is played at the same time?

SOLUTION If the intensity of the sound produced by one clarinet is I_c, the intensity level of one clarinet is

$$\beta_1 = (10\ \text{dB}) \log\left(\frac{I_c}{I_{\text{th}}}\right) = 70\ \text{dB}.$$

The second clarinet doubles the intensity, so the intensity level becomes

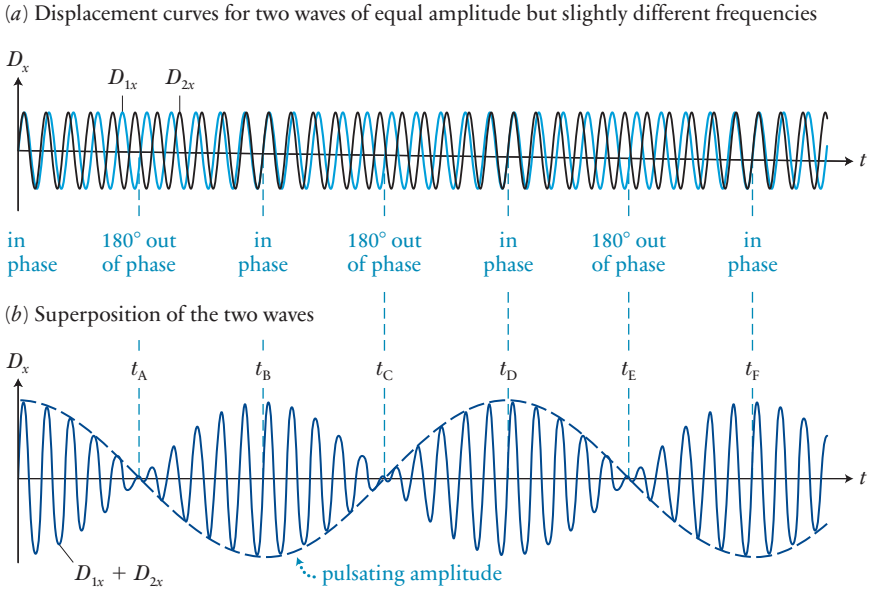
$$\beta_2 = (10\ \text{dB}) \log\left(\frac{2I_c}{I_{\text{th}}}\right) = (10\ \text{dB})\left[\log 2 + \log\left(\frac{I_c}{I_{\text{th}}}\right)\right]$$
$$= (10\ \text{dB})\log 2 + \beta_1,$$

where I have used the logarithmic relationship $\log AB = \log A + \log B$. Because $\log 2 \approx 0.3$, the intensity level increases to $\beta_2 \approx (10\ \text{dB})(0.3) + 70\ \text{dB} = 73\ \text{dB}$. So, even though the intensity doubles, the intensity *level* increases by only 3 dB. ✔

17.13 In Exercise 17.5, how many clarinets must play at the same time in order to increase the intensity level from 70 dB to 80 dB?

QUANTITATIVE TOOLS

Figure 17.27 When two sources emit waves of equal amplitude but slightly different frequencies, the waves superpose to produce a wave that has a pulsating amplitude (dashed line). We call this phenomenon *beating*.



17.6 Beats

In the first part of this chapter, the only type of interference we considered was interference between sources that emit waves of the same frequency. Let us now consider two sources that emit waves of equal amplitude but slightly different frequencies f_1 and f_2. Let the medium displacements caused by the individual waves at some fixed location be given by

$$D_{1x} = A \sin(2\pi f_1 t) \tag{17.6}$$

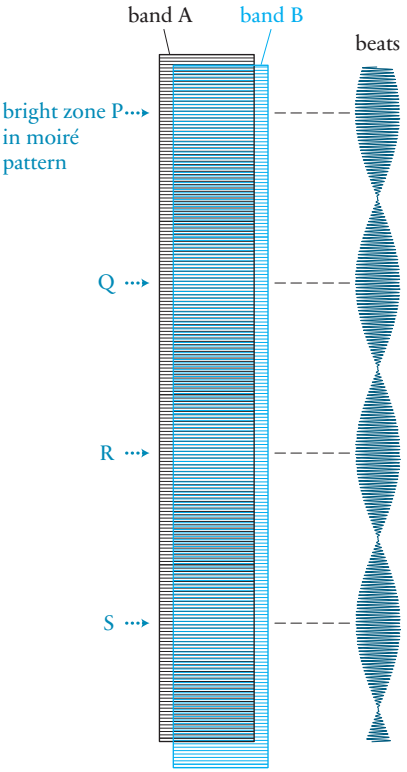
$$D_{2x} = A \sin(2\pi f_2 t). \tag{17.7}$$

These displacements are plotted in **Figure 17.27*a***. Note that because the frequencies are different, the phase relationship between the two sources is not constant. At $t = 0$, the displacements are in phase, but over time they get out of phase, with the phase of one wave getting ahead of the other. At some instants the two waves are in phase ($0, t_B, t_D, \ldots$), and at these instants their crests coincide, giving rise to oscillations of double amplitude. At other instants, however, the waves are 180° out of phase ($t_A, t_C, \ldots$), and at these instants the crests and troughs of the two waves cancel each other.

The medium displacement that results from the superposition of the two waves is shown in Figure 17.27*b*. The resultant wave is an oscillation that changes amplitude with time. At instants 0, t_B, t_D, and t_F the amplitude is maximal; at instants t_A, t_C, and t_E the amplitude is zero. This phenomenon, called **beating**, can be heard clearly with two tuning forks that oscillate at slightly different frequencies. Each maximum in the amplitude (at $0, t_B, t_D, \ldots$) corresponds to what sounds like a *beat*. When both tuning forks simultaneously produce sound, you hear a tone whose intensity periodically fades in and out.

This phenomenon is similar to the moiré pattern in **Figure 17.28**, produced by two striped bands that have slightly different line spacings. At some places the lines of the two bands overlap, producing bright regions (labeled P, Q, R, S), while at others the lines fall precisely in between one another, producing dark regions.

Figure 17.28 Moiré pattern showing beating between two patterns with slightly different line spacings.



17.14 The lines in band A in Figure 17.28 are more closely spaced than those in band B. (*a*) How many more lines does band A have between points P and S? (*b*) Describe what happens to the pattern if band B is displaced downward by one-half the spacing between its lines.

EXAMPLE 17.6 Lapping rounds

You and a friend are running laps around a race track. You begin together, and both of you run at constant speed, but you run faster than she does. Because the time interval T_y you take to complete one lap is smaller than the time interval T_f your friend takes to complete one lap, you finish the first lap ahead of her. Each lap you get farther ahead, until eventually you overtake your friend after she has run n laps (Figure 17.29). What is the frequency $f_{overtake}$ at which you overtake her? Express your answer in two ways: first in terms of T_y and T_f and then in terms of the frequencies f_y and f_f at which each of you completes one lap.

1 GETTING STARTED Regardless of how many laps my friend has run, when I overtake her I have completed exactly one lap more than she has. So, if she has completed n laps, I have completed $(n + 1)$ laps. At that instant, the process starts over, and I overtake her again when I have completed another $(n + 1)$ laps. Note that n need not be a whole number of laps; for example, if I first overtake my friend after 8.2 laps (she having completed 7.2 laps) and we both keep running at constant speed, the next overtaking occurs when I have completed 16.4 laps.

2 DEVISE PLAN The frequency at which I overtake my friend is the reciprocal of the time interval it takes me to overtake her. I know that she completes n laps, each of duration T_f, at the same instant I complete $(n + 1)$ laps, each lap of duration T_y. From this information I can determine n and the instant at which I first overtake her. The frequency at which I overtake her is the reciprocal of the period between that instant and the beginning of the race.

3 EXECUTE PLAN At the instant I first overtake my friend, I have run for a time interval $(n + 1)T_y$ and she has run for a time interval nT_f. Because we started at the same instant, I have $nT_f = (n + 1)T_y$, giving me

$$n = \frac{T_y}{T_f - T_y}.$$

The instant at which the first overtake occurs is

$$t_{overtake} = nT_f = \frac{T_y T_f}{T_f - T_y},$$

and so the overtaking frequency is

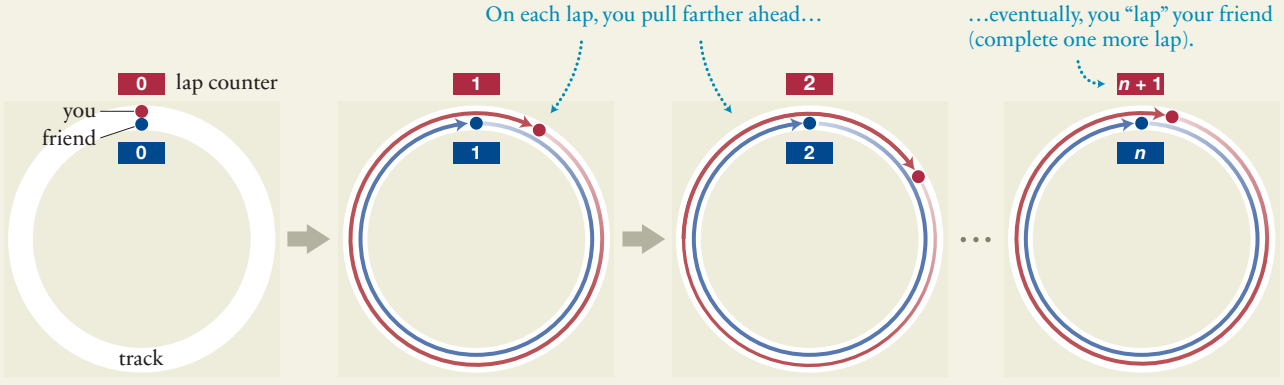
$$f_{overtake} = \frac{1}{t_{overtake}} = \frac{T_f - T_y}{T_y T_f}. \checkmark$$

To express $f_{overtake}$ in terms of f_y and f_f, I use $f_y = 1/T_y$ and $f_f = 1/T_f$ and write

$$f_{overtake} = \frac{T_f - T_y}{T_y T_f} = \frac{1}{T_y} - \frac{1}{T_f} = f_y - f_f. \checkmark$$

4 EVALUATE RESULT Because $f_y > f_f$, the overtaking frequency is positive, as I expect. My result shows that if both of us run laps at the same frequency, $f_y = f_f$, the overtaking frequency is zero. That makes sense because, running at the same rate, we never overtake each other. Another limiting case is when my friend remains at rest. In that case f_f is zero and so the overtaking frequency is equal to the frequency at which I complete laps. My answer makes sense in the limiting cases, which gives me confidence it is correct.

Figure 17.29 Example 17.6.



Suppose that in Example 17.6 your friend had run faster than you. In that case we could apply the same reasoning, but the indices y and f would now have been switched and the end result would have been $f_{overtake} = f_f - f_y$. (Because $f_y < f_f$, the overtaking frequency is again positive.) The general result, in case we do not know which frequency is greater, is thus $f_{overtake} = |f_f - f_y|$.

We can apply the reasoning I used in solving this example to the beating of two waves: The phases of the two waves catch up each time the wave of higher frequency has completed one more full oscillation than the other wave. If the wave frequencies are f_1 and f_2, then the frequency at which the beats occur, called the **beat frequency,** must, in analogy to the result from Example 17.6, be

$$f_{\text{beat}} \equiv |f_1 - f_2|. \tag{17.8}$$

We can reach the same conclusion by adding the expressions for the displacements of the two waves given by Eqs. 17.6 and 17.7:

$$D_x = D_{1x} + D_{2x} = A(\sin 2\pi f_1 t + \sin 2\pi f_2 t). \tag{17.9}$$

The sum of two sine functions can be written using the identity

$$\sin\alpha + \sin\beta = 2\cos\tfrac{1}{2}(\alpha - \beta)\sin\tfrac{1}{2}(\alpha + \beta), \tag{17.10}$$

so

$$D_x = 2A\cos\tfrac{1}{2}[2\pi(f_1 - f_2)t]\sin\tfrac{1}{2}[2\pi(f_1 + f_2)t]. \tag{17.11}$$

Because $\cos\theta = \cos(-\theta)$, we can replace the term $(f_1 - f_2)$ in Eq. 17.11 by $(f_2 - f_1)$ when $f_1 < f_2$. If we write $\Delta f = |f_1 - f_2|$ for the absolute value of the difference in frequency and $f_{\text{av}} = \frac{1}{2}(f_1 + f_2)$ for the average frequency, we can rewrite Eq. 17.11 in the form

$$D_x = 2A\cos\left[2\pi\left(\tfrac{1}{2}\Delta f\right)t\right]\sin(2\pi f_{\text{av}}t). \tag{17.12}$$

This expression tells us that the resulting wave has a frequency that is the average of the two original frequencies and an amplitude that varies with time. The factor $2A\cos\left[2\pi(\frac{1}{2}\Delta f)t\right]$ corresponds to the dashed curve in Figure 17.27*b*. It represents the slow variation in the amplitude of the combined wave.

The frequency of the amplitude variation is $\frac{1}{2}\Delta f = \frac{1}{2}|f_1 - f_2|$. However, *two* beats occur in each cycle of this amplitude variation. In Figure 17.27, one cycle in the amplitude variation lasts from $t = 0$ to $t = t_{\text{D}}$ (look at the dashed curve in Figure 17.27*b* and notice how it returns to its starting value at t_{D}). During this period *two* beats occur, one at t_{A} and the other at t_{C}. The beat frequency is therefore *twice* the amplitude variation frequency $\frac{1}{2}|f_1 - f_2|$, and so we see that the beat frequency is indeed equal to the absolute value of the difference in the frequencies of the two waves: $f_{\text{beat}} = |f_1 - f_2|$.

EXERCISE 17.7 Tuning a piano

Your middle-C tuning fork oscillates at 261.6 Hz. When you play the middle-C key on your piano together with the tuning fork, you hear 15 beats in 10 s. What are the possible frequencies emitted by this key?

SOLUTION The beat frequency—the number of beats per second—is equal to the difference between the two frequencies (Eq. 17.8). I am given the frequency of the tuning fork, $f_{\text{t}} = 261.6$ Hz, and the beat frequency, $f_{\text{B}} = (15\text{ beats})/(10\text{ s}) = 1.5$ Hz. I do not know, however, whether the frequency f_{p} of the struck middle-C piano key is higher or lower than that of the tuning fork.

If it is higher, I have $f_{\text{B}} = f_{\text{p}} - f_{\text{t}}$.

If it is lower, then $f_{\text{B}} = f_{\text{t}} - f_{\text{p}}$.

So

$$f_{\text{p}} = f_{\text{t}} \pm f_{\text{B}} = 261.6\text{ Hz} \pm 1.5\text{ Hz}$$

and the possible frequencies emitted by the out-of-tune middle-C key are 260.1 Hz and 263.1 Hz. ✔



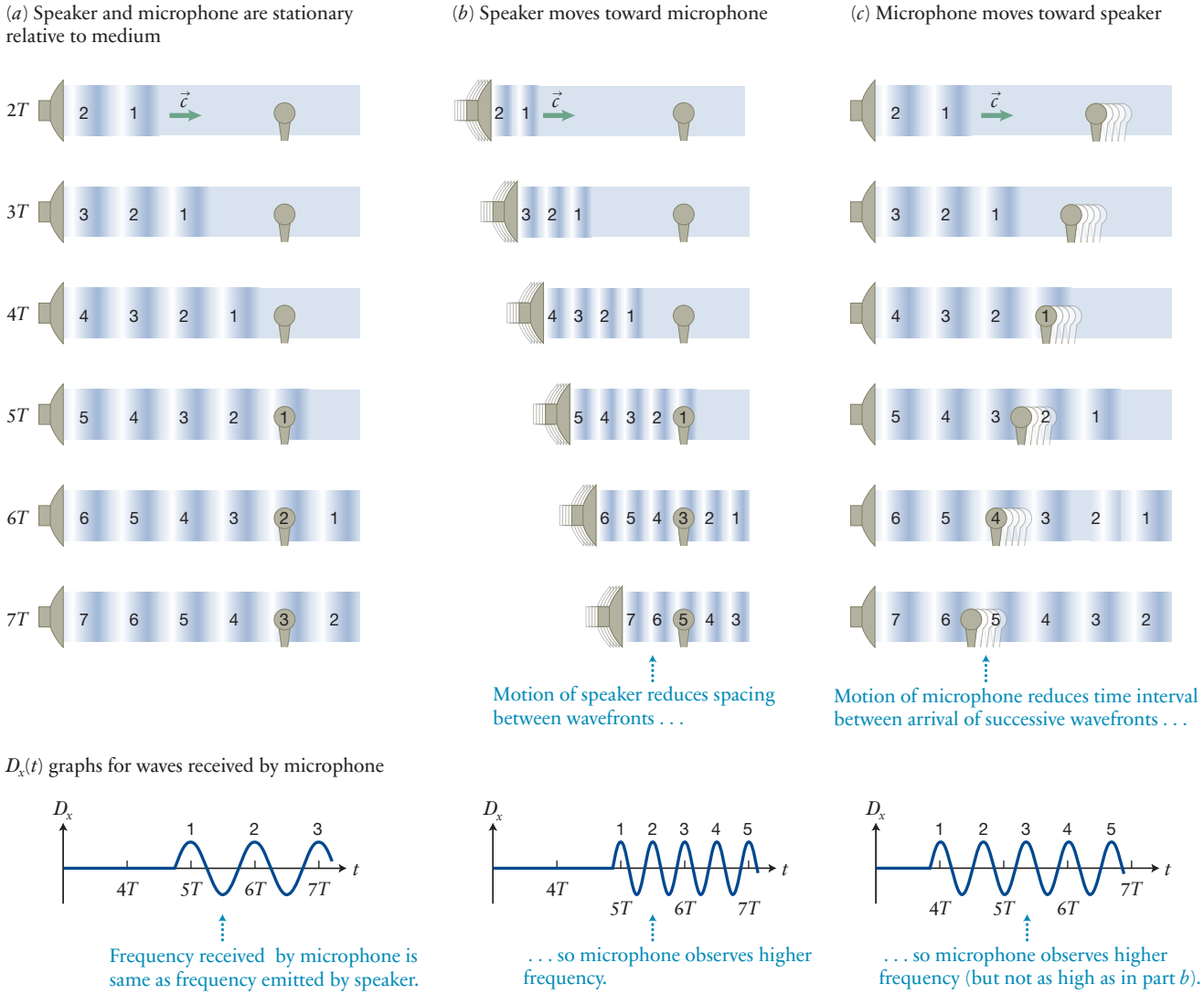
17.15 Do two sound waves of slightly different frequencies and *different* amplitudes cause beats?

17.7 Doppler effect

When an observer and a source of waves move relative to each other, the observed wave frequency is not the same as the frequency emitted by the source. When observer and source move toward each other, the observed frequency is higher than the emitted frequency; when they move away from each other, the observed frequency is lower than the emitted frequency. This phenomenon is called the **Doppler effect.**

Let us determine how the observed frequency deviates from the actual frequency of the source, depending on the speeds of source and observer relative to the medium through which the waves propagate. In **Figure 17.30*a***, a loudspeaker that is stationary relative to a medium emits a sound wave that is detected by a microphone that is also stationary relative to the medium. The period of the wave is T, and so one wavefront of a given phase (a crest, for example) leaves the loudspeaker each period T. Once the first wavefront reaches the microphone, it thereafter receives one wavefront each period T, and so the frequency received by the microphone is identical

Figure 17.30 The Doppler effect demonstrated by a speaker and a microphone. Notice that motion of the speaker toward the microphone and motion of the microphone toward the speaker both cause the frequency received by the microphone to increase—but not by the same amount. The two situations are not symmetrical because motion of the speaker shortens the wavelength of the propagating sound, whereas motion of the microphone does not.



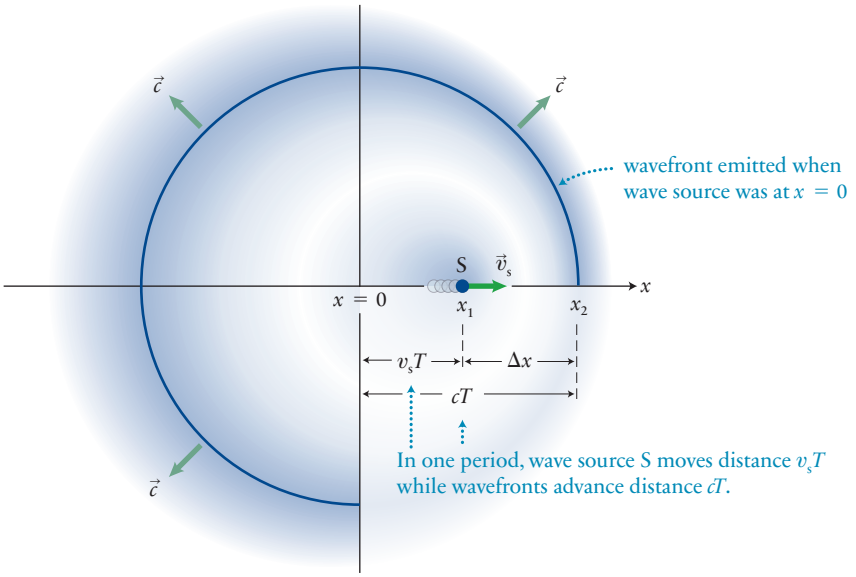
QUANTITATIVE TOOLS

to the frequency emitted by the loudspeaker. The graph at the bottom of Figure 17.30*a* shows the longitudinal medium displacement caused by the sound wave at the location of the microphone as a function of time.

Figure 17.30*b* shows what happens when the loudspeaker emits the same sound wave while moving at constant velocity toward the microphone. Because of the motion of the loudspeaker, each wavefront is emitted closer to the previous wavefront than if the loudspeaker were stationary. Because the speed of the wavefronts, which is determined by the properties of the medium (air in this case), does not depend on the motion of the loudspeaker, the more closely spaced wavefronts generated by the moving loudspeaker reach the microphone at shorter time intervals than those of the stationary loudspeaker. For example, between $t = 5T$ and $t = 7T$ only two cycles between wavefronts 1 and 3 pass the microphone when the source is stationary (Figure 17.30*a*), but four wave cycles between wavefronts 1 and 5 pass the microphone when the source is moving (Figure 17.30*b*). So the frequency detected by the microphone for the moving source is $(4 \text{ cycles})/(7T - 5T) = (4 \text{ cycles})/(2T)$, or two cycles per period instead of one cycle per period; in other words, the frequency heard by the microphone is twice the frequency emitted by the moving loudspeaker.

To obtain a general expression for the frequency received from a source S moving at speed v_s, let us evaluate the spacing between wavefronts. **Figure 17.31** shows the motion of the source during one period. (For now we consider the case when the source moves more slowly than the wave, $v_s < c$; in the next section we consider what happens when $v_s > c$.) The labeled wavefront was emitted when the source was at $x = 0$. During the time interval T, the source has moved from $x = 0$ to $x_1 = +v_sT$, at which point it emits the next wavefront. The first wavefront meanwhile has moved from $x = 0$ out to a radius $x_2 = cT$. To an observer somewhere to the right of Figure 17.31, the distance between the two wavefronts is decreased: $\Delta x = x_2 - x_1 = cT - v_sT = (c - v_s)T$. The period T_o of the wave that reaches this observer is equal to the time interval between the arrival of two

Figure 17.31 Doppler effect for a source moving toward an observer. The observer is off to the right and not shown.



successive wavefronts and is thus equal to the distance Δx divided by the speed of the waves:

$$T_o = \frac{\Delta x}{c} = \frac{c - v_s}{c} T \quad \text{(source approaching observer, } v_s < c\text{)}. \quad (17.13)$$

The frequency of the approaching source heard by the observer is thus

$$f_o \equiv \frac{1}{T_o} = \frac{c}{c - v_s} \frac{1}{T}$$

$$= \frac{c}{c - v_s} f_s \quad \text{(source approaching observer, } v_s < c\text{)}, \quad (17.14)$$

where f_s is the frequency emitted by the source. Note that because v_s is positive, $[c/(c - v_s)] > 1$, so the observed frequency is higher than the emitted frequency, as expected.

Now let us consider an observer on the opposite side of Figure 17.31. To this observer, the source is receding and the distance between wavefronts is increased: $\Delta x = (c + v_s)T$. Consequently the minus sign in front of v_s becomes a plus in Eqs. 17.13 and 17.14:

$$f_o \equiv \frac{1}{T_o} = \frac{c}{c + v_s T} \frac{1}{} = \frac{c}{c + v_s} f_s \quad \begin{matrix}\text{(source receding} \\ \text{from observer)}\end{matrix} \quad (17.15)$$

and the observed frequency is now lower than the emitted frequency, as we expect. Contrary to the situation described by Eq. 17.14, there is no limit on the speed of the source in this case.

EXAMPLE 17.8 Moving train

Standing alongside a straight section of railroad track as a train passes, you record the sound of the train's horn. By analyzing the recording, you determine that the frequency of the sound was 483 Hz as the train approached and 405 Hz after it passed. At what speed was the train moving? Use 343 m/s for the speed of sound in air. Assume that the train starts out coming almost straight at you and ends up moving almost directly away from you.

❶ **GETTING STARTED** I am given two frequencies for one sound source, so I know this problem is about a Doppler shift. I assume the train moved at constant velocity and denote its unknown speed by v. Because the horn—that is, the sound source—was on the train, v is also the speed at which the horn moved, and I know I can use some Doppler equation to calculate this speed. The wave speed is given as $c = 343$ m/s.

❷ **DEVISE PLAN** As the train approaches, the frequency I hear is given by Eq. 17.14. After it passes, the frequency is given by Eq. 17.15. This gives me two equations with two unknowns, f_s (the frequency of the source) and v, so I can solve for v.

❸ **EXECUTE PLAN** As the train approaches, the observed frequency is given by Eq. 17.14:

$$f_1 = \frac{c}{c - v_s} f_s = \frac{c}{c - v} f_s. \quad (1)$$

After the train passes, the observed frequency is given by Eq. 17.15:

$$f_2 = \frac{c}{c + v_s} f_s = \frac{c}{c + v} f_s. \quad (2)$$

Solving Eq. 2 for f_s and substituting the result into Eq. 1, I obtain

$$f_1 = \frac{c}{c - v}\left(\frac{c + v}{c} f_2\right) = \frac{c + v}{c - v} f_2.$$

Multiplying both sides by $(c - v)$ gives

$$f_1(c - v) = f_2(c + v)$$

$$v(f_1 + f_2) = c(f_1 - f_2).$$

Therefore

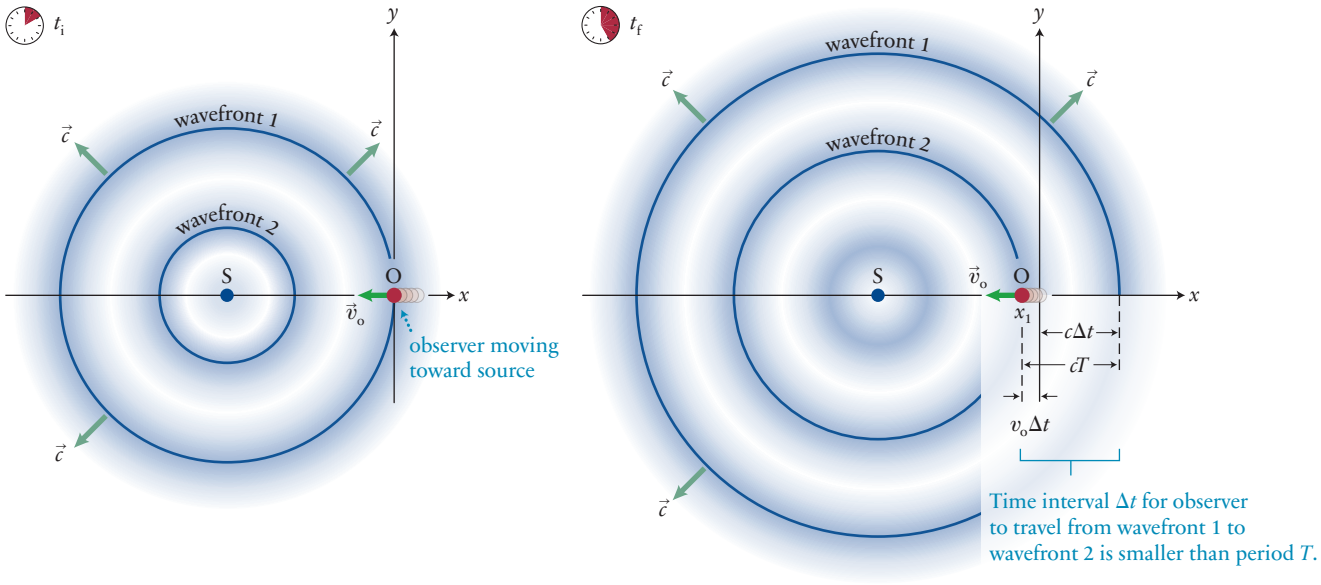
$$v = \frac{f_1 - f_2}{f_1 + f_2} c = \frac{483 \text{ Hz} - 405 \text{ Hz}}{483 \text{ Hz} + 405 \text{ Hz}} (343 \text{ m/s}) = 30 \text{ m/s}. \checkmark$$

❹ **EVALUATE RESULT** This is about 110 km/h or 70 mi/h, not unreasonable for a train.

17.16 In Example 17.8, what is the frequency of the sound the horn makes?

QUANTITATIVE TOOLS

Figure 17.32 Doppler effect for an observer moving toward the source.



Applications of the Doppler effect range from measuring the speed of blood flow by reflecting a beam of ultrasound waves from red blood cells to measuring the speed of water molecules in a storm using radar. The latter uses electromagnetic waves (see Chapter 30), which, like sound waves, exhibit a Doppler effect.

Any movement by an observer also affects the frequency of sound waves. When an observer is moving toward a stationary source, as in Figure 17.30*c*, the observer travels through the oncoming wavefronts and consequently observes more of them per unit time than if she or he were at rest. Between $t = 4T$ and $t = 6T$ in Figure 17.30*c*, for example, three wave cycles pass the microphone, and so the observed frequency is $(3 \text{ cycles})/(6T - 4T) = (1.5 \text{ cycles})/T = 1.5f_s$. Because the microphone moves toward the loudspeaker, each subsequent wavefront has to travel a shorter distance to reach the microphone, and so the observed frequency is higher than the emitted frequency.

Figure 17.32 allows us to relate the increase in frequency to the speed v_o of an observer O who is moving toward the source S. At t_i the observer is at $x = 0$ just as wavefront 1 passes her. At t_f the observer meets wavefront 2 at position x_1, and so she detects one complete cycle (that is, two wavefronts) in $\Delta t = t_f - t_i$. During that time interval, wavefronts 1 and 2 have moved a distance $c\Delta t$. From the figure we see that the sum of the distances traveled by the wavefront and the moving observer is equal to the distance between wavefronts, so $c\Delta t + v_o\Delta t = cT$. The time interval at which successive wavefronts reach the observer is thus

$$\Delta t = \frac{cT}{c + v_o}. \tag{17.16}$$

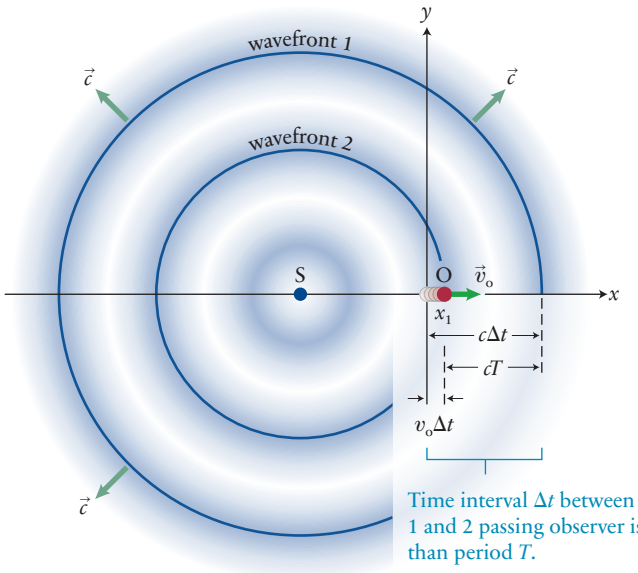
The frequency she observes is

$$f_o \equiv \frac{1}{\Delta t} = \frac{c + v_o}{c}\frac{1}{T} = \frac{c + v_o}{c}f_s \quad \text{(observer approaching source)}. \tag{17.17}$$

Because $(c + v_o) > c$, the observed frequency is higher, as expected.

Let us now consider an observer who moves away from the source at a speed that is smaller than the speed of sound, $v_o < c$. (If the observer were

Figure 17.33 Doppler effect for an observer moving away from the source at a speed $v_o < c$.



to move away at a greater speed, the sound would never catch up with the observer.) Again, we take the observer to be at $x = 0$ just as wavefront 1 passes him at t_i. **Figure 17.33** shows the situation at t_f when wavefront 2 catches up with the observer at position x_1. In the time interval $\Delta t = t_f - t_i$ the observer hears a complete cycle. During that time interval, wavefronts 1 and 2 have moved a distance $c\Delta t$. From the figure we see that the distance between wavefronts is equal to the difference between the distances traveled by the wavefront and the moving observer, so $c\Delta t - v_o\Delta t = cT$. The time interval at which successive wavefronts reach the observer is thus

$$\Delta t = \frac{cT}{c - v_o}. \tag{17.18}$$

Consequently the plus sign in front of v_o in Eq. 17.17 becomes a minus, and the observed frequency is thus

$$f_o \equiv \frac{1}{\Delta t} = \frac{c - v_o}{c}\frac{1}{T}$$

$$= \frac{c - v_o}{c} f_s \quad \text{(observer receding from source, } v_o < c\text{)}. \tag{17.19}$$

Because $(c - v_o) < c$, the observed frequency is lower, as expected.

If both source and observer are moving, the two effects multiply—the moving source causes wavefronts to be spaced differently, giving rise to the observed frequency given in Eqs. 17.14 and 17.15, and the moving observer detects the wavefronts at a different rate, as indicated by Eqs. 17.17 and 17.19. The two effects can be combined in one equation:

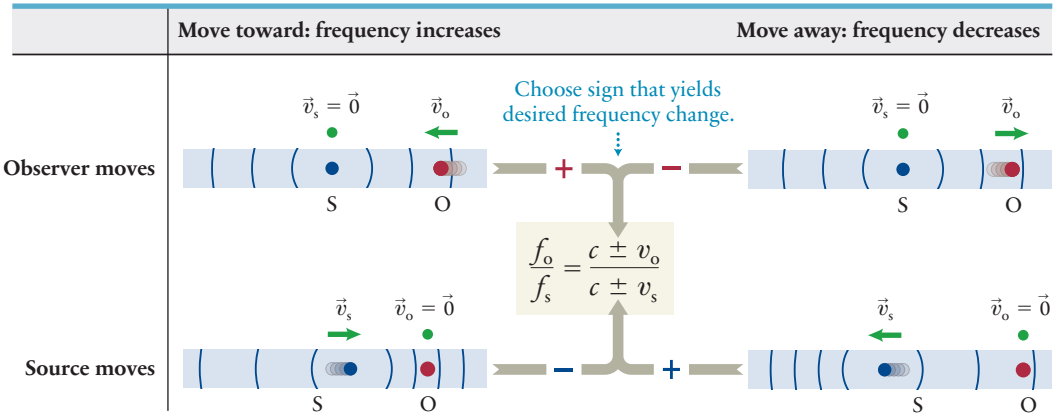
$$f_o = \left(\frac{c \pm v_o}{c}\right)\left(\frac{c}{c \pm v_s}\right) f_s = \frac{c \pm v_o}{c \pm v_s} f_s, \tag{17.20}$$

which can be written in the more symmetrical form

$$\frac{f_o}{f_s} = \frac{c \pm v_o}{c \pm v_s}. \tag{17.21}$$

QUANTITATIVE TOOLS

Figure 17.34 To use the general Doppler equation, the key thing you need to remember is that the observed frequency increases when source and observer move toward each other and decreases when they move apart.



Keep in mind that in Eqs. 17.20 and 17.21 the speeds of the observer and source are measured relative to the medium. Also, these expressions are subject to the same conditions as in Eqs. 17.14 and 17.19: when the source is approaching the observer $v_s < c$, and when the observer is receding from the source $v_o < c$.

Rather than trying to remember the correct choices for the signs in front of v_o and v_s in this expression, it is easiest to remember that the observed frequency increases when source and observer move toward each other and decreases when they move apart. For example, the top left of **Figure 17.34** shows an observer approaching a stationary source ($v_s = 0$). The observed frequency is higher and therefore you select the plus sign in front of v_o in Eq. 17.21. When the observer moves away from the stationary source (top right of Figure 17.34), the observed frequency is lower and so you select the minus sign in front of v_o in Eq. 17.21. The bottom left of Figure 17.34 shows a source moving toward an observer who is at rest relative to the medium ($v_o = 0$). The observed frequency is higher, and therefore you select the minus sign in front of v_s in Eq. 17.21. In the bottom right of Figure 17.34 the source moves away from the stationary observer; the observed frequency is now lower and you select the plus sign in front of v_s in Eq. 17.21.

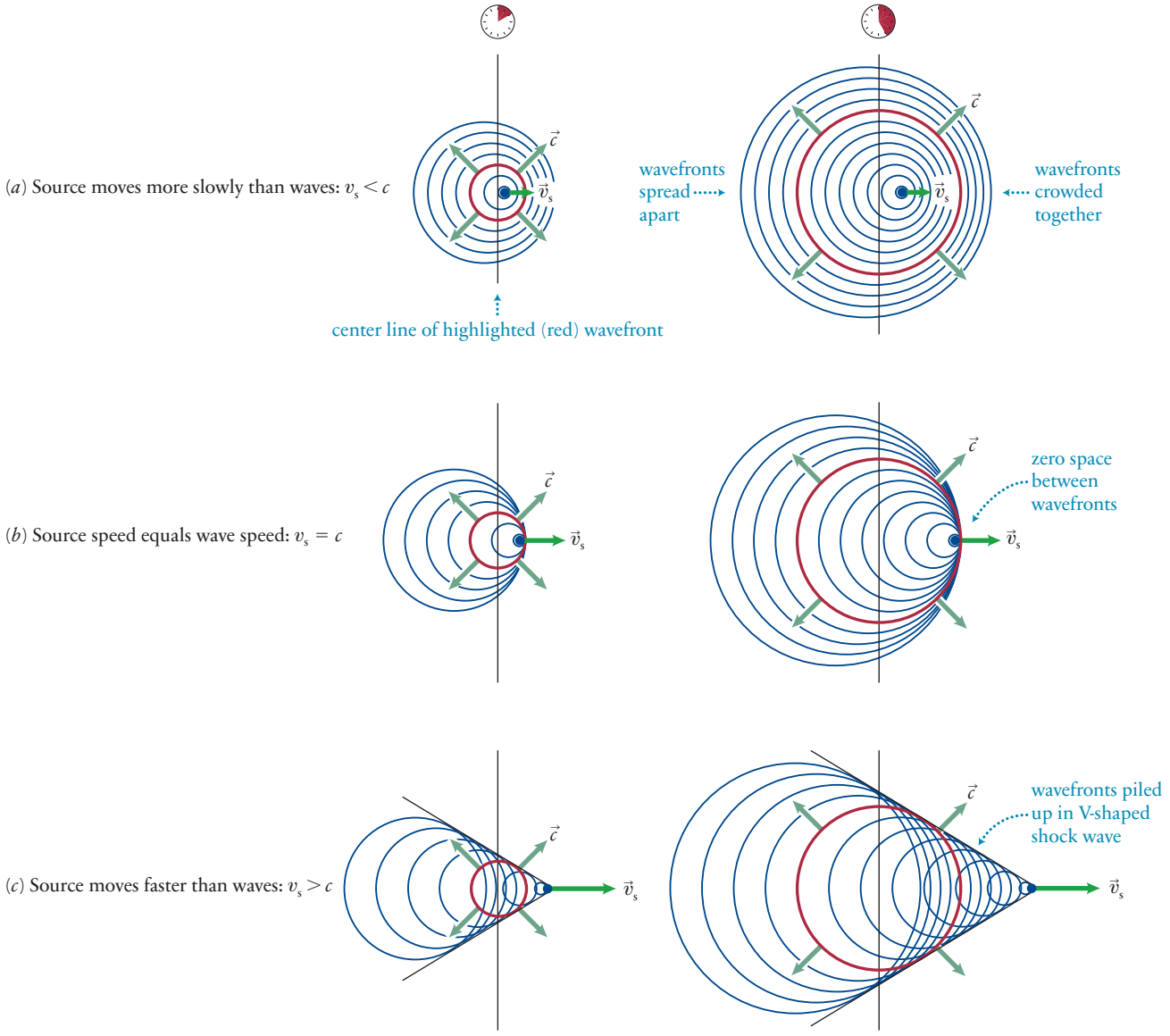


17.17 (*a*) Suppose an observer moves at half the speed of sound toward a stationary source of sound waves. By what factor does the observed frequency differ from the emitted frequency? (*b*) By what factor does the observed frequency differ if, instead of the observer moving, the source moves toward the stationary observer at half the speed of sound? (*c*) Why is there a difference?

17.8 Shock waves

The observed frequency f_o given by Eq. 17.14 becomes infinite when the speed of the source approaches the speed of sound. As can be seen in **Figure 17.35a** on the next page, the wavefronts in front of a source moving at a speed $v_s < c$ crowd together. When a source moves at the speed of sound, *all* wavefronts in the forward direction travel with the source, and so the spacing between them is reduced to zero. A source traveling *faster* than the speed of sound overtakes the wavefronts it has generated. The wavefronts then pile up to form a wavefront that is wedge-shaped in two dimensions and cone-shaped in three dimensions.

Figure 17.35 The formation of a shock wave as the speed of a source reaches and then exceeds the speed of sound in the medium.



This wedge- or cone-shaped wavefront, called a **shock wave,** occurs whenever the speed of any object moving through a medium is higher than the the speed of sound in the medium, *even if the object emits no sound*. A bullet moving faster than the speed of sound in air, for example, creates a shock wave (**Figure 17.36**) even though the bullet emits no sound. Air piles up in front of the speeding bullet like snow in front of a snowplow, and this piling up of the air creates a compression in front of the bullet. This compression then moves outward at the speed of sound, just as the compression in Figure 17.8 does, and forms the wedge- or cone-shaped shock wave, which trails the bullet like a water wake trailing behind a ship.

When a shock wave sweeps over your ears, the piled-up wavefronts sound like an explosion, just as a small boat gets jolted when struck by the wake of a passing speedboat. For a small object, such as a bullet, the shock wave produces a sharp, cracking sound. The crack of a long whip is a shock wave produced when the tip of the whip is made to move faster than the speed of sound. For large objects, the shock wave sounds like a thunderous boom called a *sonic boom*. The sonic boom of a supersonic

Figure 17.36 Shock wave produced by a bullet traveling through air.



airplane (a plane with a speed higher than the speed of sound) reaches the pain threshold of the human ear even at a distance of 20 km. For this reason, supersonic transport is not feasible over land. The cone-shaped shock wave dragging behind the plane would break windows on the ground even at its maximum flying altitude of about 20 km.

17.18 Does an observer moving at twice the speed of sound toward a stationary sound source detect a sonic boom?

As can be seen in **Figure 17.37**, the higher the source speed, the narrower the wedge or cone formed by the shock wave. From the right-angle triangle PQR, we see that the angle the shock wave makes with the direction of motion is given by

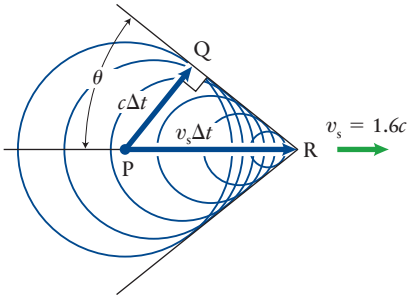
$$\sin\theta = \frac{c}{v_s} \quad (v_s > c). \tag{17.22}$$

For motion through air, the ratio v_s/c is called the *Mach number*. A plane flying at Mach 1.8, for example, has a speed of 1.8 times the speed of sound.

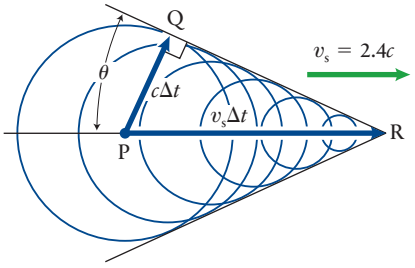
17.19 Estimate the speed of the bullet in Figure 17.36.

Figure 17.37 The shape of a shock wave depends on the speed of the source. The higher the source speed, the smaller the angle θ and so the sharper the cone or wedge.

(*a*) Lower source speed v_s: wider angle θ



(*b*) Higher source speed v_s: narrower angle θ



Chapter Summary

Characteristics of waves in two and three dimensions (Sections 17.1, 17.4, 17.5)

Concepts A **wavefront** is a curve or surface in a medium on which all points of a propagating wave have the same phase.

A **planar wavefront** is a flat wavefront that is either a plane or a straight line.

A **surface wave** is a wave that propagates in two dimensions and has circular wavefronts. A **spherical wave** is a wave that propagates in three dimensions and has spherical wavefronts.

According to **Huygens' principle,** any wavefront may be regarded as a collection of many closely spaced, coherent point sources.

Diffraction is the spreading out of waves either around an obstacle or beyond the edges of an aperture. The effect is more pronounced when the size of the obstacle or aperture is about equal to or smaller than the wavelength of the wave.

Quantitative Tools If no energy is dissipated, the amplitude A of a wave originating at a point source decreases with increasing distance r from the source as

$$A \propto \frac{1}{\sqrt{r}} \quad \text{(surface wave)}$$

or

$$A \propto \frac{1}{r} \quad \text{(spherical wave)}.$$

The **intensity** I (in W/m^2) of a spherical wave that delivers power P to an area A oriented normal to the direction of propagation is

$$I \equiv \frac{P}{A}. \tag{17.1}$$

If a point source emits waves uniformly in all directions at a power P_s and no energy is dissipated, the intensity a distance r from the source is

$$I = \frac{P_s}{4\pi r^2}. \tag{17.2}$$

The intensity I_{surf} (in W/m) of a surface wave that delivers power P to a length L oriented normal to the direction of propagation is

$$I_{surf} \equiv \frac{P}{L}. \tag{17.3}$$

Sound waves (Sections 17.2, 17.5)

Concepts **Sound** is a longitudinal *compressional* wave propagating through a solid, liquid, or gas. The wave consists of an alternating series of *compressions* (where the molecules of the medium are the molecules are spaced far apart). The frequency range of *audible* sound is 20 Hz to 20 kHz.

The speed of sound c depends on the density and elastic properties of the medium. In dry air at 20 °C, the speed of sound is 343 m/s.

Quantitative Tools The *threshold of hearing* I_{th} is the minimum sound intensity audible to humans. For a 1.0-kHz sound,

$$I_{th} \approx 10^{-12}\ \text{W/m}^2.$$

For a sound of intensity I, the **intensity level** β in decibels is

$$\beta \equiv (10\ \text{dB}) \log\left(\frac{I}{I_{th}}\right). \tag{17.5}$$

Interference effects (Sections 17.3, 17.6)

Concepts Two or more sources emitting waves that have a constant phase difference are called **coherent sources.** If that constant phase difference is zero, the sources are said to be *in phase*.

Along **nodal lines,** waves cancel each other, and so the displacement of the medium is zero. Along **antinodal lines,** the displacement of the medium is a maximum.

The superposition of two waves of equal amplitude but slightly different frequencies results in a wave of oscillating amplitude. This effect is called **beating.**

Quantitative Tools When two waves of frequencies f_1 and f_2 result in **beating,** the **beat frequency** is

$$f_{beat} \equiv |f_1 - f_2|, \tag{17.8}$$

and the displacement D_x of the particles of the medium is

$$D_x = 2A \cos\left[2\pi\left(\tfrac{1}{2}\Delta f\right)t\right] \sin(2\pi f_{av} t), \tag{17.12}$$

where $\Delta f = |f_1 - f_2|$ and $f_{av} = \frac{1}{2}(f_1 + f_2)$.

The effects of motion on sound (Sections 17.7, 17.8)

Concepts The **Doppler effect** is a change in the observed wave frequency caused by the relative motion of a wave source and an observer.

A **shock wave** is a conical (wedge-shaped) disturbance caused by the piling up of wavefronts from a source moving at a speed greater than or equal to the wave speed in the medium.

Quantitative Tools If a source moving with speed v_s relative to the medium produces sound of frequency f_s, the **Doppler effect** causes an observer moving with speed v_o relative to the medium to observe the sound as having a frequency f_o given by

$$\frac{f_o}{f_s} = \frac{c \pm v_o}{c \pm v_s}. \tag{17.21}$$

The $\pm$ signs are chosen so that $f_o > f_s$ when the source and observer approach each other and $f_o < f_s$ when they move apart.

As a shock wave propagates at speed c, the angle θ it makes with the direction in which the source moves is given by

$$\sin\theta = \frac{c}{v_s} \quad (v_s > c), \tag{17.22}$$

where v_s is the speed of the source relative to the medium. The ratio v_s/c is the *Mach number*.

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

17.1 Wavefronts

1. A neat trick for determining whether or not a train is coming is to press your ear to a railroad track. Why is this much more effective than listening for the train's sound carried by air? •
2. Suppose a material is stiffer in the x direction than in the y direction. Draw the wavefronts emanating from a point source embedded in this material and radiating in these two directions. ••
3. Your professor delivers a lecture. Does the amplitude of sound waves in the classroom decrease like $1/r$, where r is the distance from the professor to you? ••
4. A cannon is fired some distance away from you, and you wish to estimate that distance by determining how much sound energy enters your ear. (*a*) How does the sound energy depend on the distance between the cannon and your ear? (*b*) Now you use the dependence you described in part *a* to judge how far away the cannon is. First you estimate the distance assuming that none of the energy in the sound waves is dissipated (into, say, random motion of the molecules that make up the air), and then you repeat the estimate assuming that some of the sound energy is dissipated. Is your first estimate higher or lower than your second estimate? ••
5. One form of whispering gallery is an elliptical room where a person standing at one focus can hear quite clearly someone speaking very quietly at the other focus. Use your knowledge of ellipses to explain how this works. •••

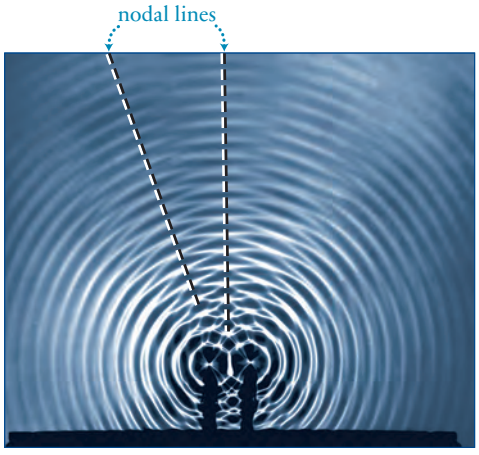
17.2 Sound

6. Describe the pressure of particles at points in a sound wave where the gas is (*a*) compressed at maximum, and (*b*) minimally compressed. •
7. Because the audible range in bats is about 10 kHz to more than 120 kHz, these animals are oblivious to normal human conversation, which is at frequencies of typically several hundred hertz. If you shout, though, you may in fact startle a bat. Why? ••
8. Standing waves (see Section 16.6) can form when sound traveling inside a closed tube reflects from each end. These waves have a node at each (immovable) end. Standing waves can also form in a tube with open ends. In this case, the waves have antinodes rather than nodes at the ends. Use the requirement of antinodes at the tube ends to determine the relationship between wavelength and tube length for the fundamental and the next two harmonics for standing waves in a tube that is open at both ends. ••
9. Sketch the three lowest-frequency harmonic wave function patterns for standing waves in the air in a tube that is open at both ends, (*a*) representing *displacement* nodes and antinodes and (*b*) representing *pressure* nodes and antinodes. Your sketches should be similar to the pattern in Figure 17.9*c*. ••

17.3 Interference

10. If the frequency of the sources in Figure P17.10 is increased, do the nodal lines get farther apart, get closer together, or remain unchanged? •

Figure P17.10

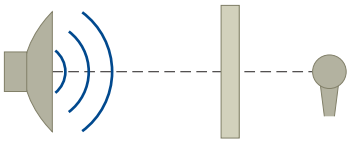


11. What happens to the nodal lines in Figure P17.10 if the phase of one of the sources is shifted by 180°? •
12. Are stable nodal lines formed by two sources that emit waves that have the same amplitude but different frequencies? ••
13. Speaker manufacturers often discourage placing speakers in front of curtains or drapes because having an uncovered wall behind the speakers often alleviates dead listening spots in the room. Explain why a wall helps. ••
14. How many nodal lines are created by two point sources that are separated by a distance equal to 2.5λ? ••
15. Vibration in a car can create concentric ripples on the surface of a cup of coffee sitting on the dashboard. What kind of waves are the ripples, and why are they circular? ••
16. Two sound sources emitting at the same wavelength λ are placed a distance d_s apart and a perpendicular distance d_m away from a row of microphones. Show that the distances Δx between those microphones that experience maximum variations in loudness are given by $\Delta x = d_m\lambda/d_s$ in the limit $d_m \gg d_s$. •••
17. Two coherent sources of water waves of frequency f and speed c are a distance d apart and equal in amplitude. Nodal lines exist where the path lengths from the sources differ by an odd number of half wavelengths (where the contributions are 180° out of phase). (*a*) Far from the sources (where we can treat wavefronts from each source as parallel), what are the angles between the nodal lines in the water and the perpendicular to the line joining the sources? (*b*) If you introduce a phase difference ϕ between the sources, how do the directions of the nodal lines shift? •••
18. Three equally spaced, coherent sources of water waves of frequency f are in a line with a distance d separating each adjacent pair. All three sources have the same amplitude. Nodal lines exist where the separate waves from the three sources cancel each other. Far from the sources, where we can treat the wavefronts from each source as parallel, what are the angles between the nodal lines in the water and the perpendicular to the line joining the sources? •••

17.4 Diffraction

19. When your mother shouted at you to come inside when you were a kid, how could you hear her when you were around a corner? •
20. The speaker in Figure P17.20 faces a narrow obstructing panel made of sound-absorbing material. A microphone is placed dead center behind the panel. Does the microphone pick up sound or not? ••

Figure P17.20



21. Two very small, coherent sources are separated by $d = 2\lambda$, where λ is the wavelength of the circular waves created by each source. (*a*) Use Huygens' principle to construct a dozen wavefronts from each source, six troughs and six crests. Mark antinodes with blue lines and mark nodes with red lines, starting each line at least two wavelengths from either source. (*b*) Repeat for the case when the two sources are not coherent, so that waves from one source are exactly one-half wavelength out of phase with waves from the other. (*c*) Compare your diagrams. Are the original antinodes still antinodes? Are the original nodes still nodes? ••
22. Phasors (see Section 15.5) are useful for adding waves that are out of phase. For example, assume you have two sources, 1 and 2, that emit sound waves of the same frequency f. At your detector, the waves from source 1 have an amplitude of 1.00×10^{-8} m, and those from source 2 have an amplitude of 0.600×10^{-8} m and a phase of $+\pi/3$ relative to the phase of the source 1 waves. If the phase of the waves from source 1 is arbitrarily assigned to be $\phi_1 = 0$, what are the amplitude and phase of the combined waves? (Hint: It will help to draw the two waves as phasors.) •••

17.5 Intensity

23. You measure the sound intensity of a sound wave to be 1.50 W/m^2 at 2.60 m away from a point source. Calculate the power output of the source. •
24. Radio Station DXYZ has a power output of 128 W. While listening at home, you measure the intensity of the sound to be 3.40×10^{-6} W/m^2. How far away from your home is the radio station? ••
25. One mosquito emits a sound of intensity level 15 dB, as measured on a decibel meter 0.50 m away. What is the intensity level of the sound emitted by 100 mosquitoes at the same distance? •
26. If ten screaming children on the playground produce a sound for which the intensity level is 80 dB, how many children do you need to produce an 82-dB sound? •
27. A sound source at the center of a cubical room radiates uniformly in all directions. What is the difference in decibels between the intensity level at a ceiling corner of the room and the intensity level at the center of one of the walls? Ignore reflections. ••
28. At a distance of 3.0 m from your stereo system with two speakers, what power output must your stereo amplifier have in order for the intensity level to be 80 dB? Assume that all the energy from the amplifier goes into producing sound. ••
29. A satellite orbiting Earth at an altitude of 450 km emits a radio signal that has a power level of 12 W when it reaches Earth. The signal is beamed in such a way that it covers an area of 8.0 m^2 on Earth's surface. What power is received by a ground-based antenna if the surface area of the antenna is 400 mm^2? ••
30. At a basketball game in an enclosed arena, you happen to sit next to a screaming baby. If the baby's yell arrives at your ear with an intensity level of 75 dB, and the music that reaches your ear from the loudspeakers during the teams' warm-up is at 80 dB, (*a*) what is the intensity I of the overall sound you hear and (*b*) what is the intensity level of the sound you hear? ••
31. The intensity level from a directional loudspeaker at your position at a political rally is 95 dB. (*a*) What is the intensity I of the sound in watts per square meter? (*b*) What is the

power output of the loudspeaker if you are 20 m from the speaker and it emits sound into one hemisphere only? ••

32. You are on the design team for a new airport. The developer wants to build a hotel as close to the airport as possible. The intensity level of a large jet during takeoff is 140 dB at 50 m from the jet. The developer wants to build the hotel such that the intensity level heard by guests is not higher than 50 dB inside the hotel rooms and not higher than 125 dB in the parking lot of the hotel. (*a*) How far from the airport would you suggest the developer build the hotel? (*b*) What is the maximum allowable intensity I of the sound in the hotel rooms? ••
33. The intensity level at the front of a library reading room is 70 dB when 120 students are in the room and 20 dB when the room is empty. What do you expect the intensity level to be when 60 students are in the room? ••
34. A typical car horn has an intensity level of 90 dB at a distance of 1.0 m. (*a*) What is the intensity level of 16 cars all honking their horns at an intersection on a city street? (*b*) What is the power output of the combined horns at a distance of 5.0 m? (*c*) What is the intensity level three city blocks away? Assume 100 m for the length of each block. ••
35. On a whale-watching expedition, your underwater sound detector picks up a whale sound that has an intensity of 9.0 μW/m^2. (*a*) What is the intensity level β of this sound? (*b*) If your detector tells you that you are 2.3 km from the whale, what is the power of the emitted signal? (*c*) How many whales, emitting sound at the same intensity $I = 9.0\ \mu$W/m^2, are needed to produce an intensity level of 100 dB at your location? ••
36. A frequency generator (a device that produces sinusoidal signals whose frequency and amplitude can be independently set) is used to power a horn. The frequency is set to 2000 Hz, and for a particular amplitude the sound intensity is measured to be 0.050 W/m^2 at a distance of 12 m. (*a*) What is the power output of the horn in this case? (*b*) If the power output is increased to 92 W, what is the new intensity reading? ••
37. Suppose that two coherent, spherical sound sources, 1 and 2, each generate the same amount of power P. The distance between the sources is 2λ. Point R is on a line of maximum destructive interference, and the distance between R and source 1 is equal to 9λ. Point Q is on a line of maximum constructive interference, and the distance between Q and source 1 is 6λ. Obtain expressions for the intensity at points R and Q. •••
38. Your bedroom is 8.0 m away from the living room. When requested by your parents to turn down your stereo, you do, and the intensity level where they are sitting in the living room drops from 50 dB to 45 dB. What is the change in the power of the sound emitted by your stereo? •••

17.6 Beats

39. Examine each situation for two waves and determine if audible beats can be produced. Two wave with (*a*) the same amplitude; (*b*) the same frequency; (*c*) slightly different amplitude; and (*d*) slightly different frequency. ••
40. Two stationary tuning forks with frequencies of 246 Hz and 252 Hz are struck simultaneously. Calculate the frequency of the resulting sound. •
41. Two violin players are next to each other on a theater's stage. One plays a 350-Hz note, while the other simultaneously plays a 353-Hz note. (*a*) What is the beat frequency heard by the audience? (*b*) How many beats does the audience hear in a minute? ••
42. Two tuning forks are struck simultaneously, the first one having a frequency of 762 Hz. If a nearby listener hears 20 beats per second, determine the possible frequency or frequencies of the other tuning fork. ••
43. You have four electronic audio generators, A, B, C, and D, all producing tones near 1200 Hz that are easily audible but not distinguishable by ear. When you increase the frequency of generator A a little, the beat frequencies between A and B, A and C, and A and D all increase. When you increase the frequency of generator B a little, the B-C beat frequency increases and the B-A and B-D beat frequencies decrease. Rank the generators by frequency, from lowest to highest. ••
44. One tuning fork vibrates at a frequency of 528 Hz, and a second one vibrates at 524 Hz, but you don't know which is which. After applying a small piece of modeling clay to the end of one fork, you notice that the beat frequency has increased. On which fork did you put the clay? ••
45. Three tenors are giving a concert at an outdoor amphitheater. What are the possible beat frequencies if tenor 1 holds a 262-Hz note while tenor 2 holds a 264.3-Hz note and tenor 3 holds a 258-Hz note? ••
46. When two frequencies are added together, the beat frequency and the average frequency are an octave apart. What is the ratio of the original frequencies? ••
47. Two violin strings, each having a linear mass density of 0.0014 kg/m and under the same 100-N tension, have the same fundamental frequency of 660 Hz. If the tension in one of the strings is changed to 102 N and both strings vibrate in their fundamental nodes, what is the beat frequency? ••
48. During a fire drill in your dormitory, you notice that the sound waves from the two alarms on your floor interfere with each other; you hear five beats every two seconds. (*a*) If the average of the two frequencies is 3500 Hz, what are the frequencies of the two alarms? (*b*) If your room is 15 m from each alarm, what is the intensity I of the sound if the power output is 60 W for each alarm? ••

17.7 Doppler effect

49. Describe the pitch of an approaching siren if it is (*a*) coming close to you and (*b*) after it passes you. •
50. While driving down the road, you spot a friend standing on the sidewalk. You honk your horn at him while driving 70 km/h. Your horn emits a frequency of 360 Hz. What frequency does your friend hear (*a*) as you approach and (*b*) after you pass by? •
51. The electric bells at your school buzz with a frequency of 400 Hz. If you are late for class and riding your bike down a hill toward the school at a constant speed of 4.47 m/s, what is the frequency of the sound you hear? •
52. A stationary mechanical motor generates 8.0 Hz water waves moving at 3.0 m/s. What is the frequency of these waves as observed by a person in a motorboat that approaches the source at 12 m/s? •

53. How fast do you have to travel away from a stationary sound source in order for the frequency to be shifted by (*a*) 1%, (*b*) 10%, and (*c*) a factor of 2? ••
54. While you watch a parade, a band on a float passes you. You detect the frequency of a note played on a flute to be 352 Hz when the float is coming toward you and 347 Hz after the float passes you. At what speed is the float traveling? ••
55. You and a friend are riding bikes, you moving at 9.72 km/h and she moving at 7.20 km/h. She is behind you and turns on her electric horn, which has a frequency of 300 Hz. (*a*) What frequency do you hear? (*b*) What frequency is heard by a person who is standing alongside your route after you both have passed him? ••
56. While you are driving on a straight road at 97 km/h, a police car traveling in the opposite direction comes toward you with the siren blaring, and the frequency you hear is 310 Hz. Starting when the two of you are 200 m apart, it takes you 3.00 s to pass the police car. What is the frequency of the siren? ••
57. A car traveling west at 90 km/h along a road running parallel to a railroad track approaches a train traveling east at 65 km/h. If the train's whistle emits a frequency of 400 Hz, what is the frequency heard by the driver of the car (*a*) as the car approaches the train and (*b*) after the two vehicles have passed each other? ••
58. A foghorn that emits a frequency of 150 Hz is mounted on a buoy in the bay. If a ship approaches the buoy at 21.2 km/h, (*a*) what frequency does the captain of the ship hear? (*b*) Hearing the foghorn is a signal for the captain to slow his ship. After passing the buoy, the captain hears the horn again, but the frequency he detects is now 149 Hz. What is the new speed of the ship? ••
59. Two airplanes simulating a dogfight are headed right for each other, plane A moving at 285 km/h and plane B moving at 295 km/h. The electronic machine gun on A can emit 300 rounds per minute, with each round accompanied by a sound pulse. (*a*) How many pulses does the pilot in B hear during a burst that lasts 10 s according to observer A? (*b*) How long does the burst last for the pilot in B? ••
60. You stand in your yard and hold a battery-powered buzzer that is emitting a frequency of 560 Hz as a friend stands next to you. (*a*) What frequency does your friend hear if you throw the buzzer away from you with a speed of 18.0 m/s? (*b*) You and your friend now get on a flat railroad car that begins to travel at 12.0 m/s, and you throw the buzzer in the direction opposite the direction of the car's velocity. What frequency does your friend hear? (*c*) What frequency would a person standing alongside the tracks hear? ••
61. A dog whistle has a frequency of 21 kHz, which is above the high end of the audible range for humans. At what minimum speed and in what direction must you travel in order to just hear the whistle? ••
62. During trials of a new submarine, its speed is measured with an underwater sonar gun anchored to the ocean floor. The gun emits sound waves at a frequency f, and these waves reflect off the surface of the submarine and return to the source at a frequency $f' \neq f$. If the submarine is receding from the sonar gun at speed v, derive an expression for f' in terms of f, v, and c (the speed of the waves through water). •••
63. A device that both emits and receives sound waves is traveling at 80 km/h. The emitted waves have a frequency of 700 Hz. If the waves are reflected off a stationary flat surface in front of the device, what frequency is detected by the device? •••
64. Two students standing 54.9 m apart are tossing around footballs designed to whistle while they are airborne. (*a*) When student A throws her football toward student B at 16.0 m/s, it emits a frequency of 680 Hz. What frequency is heard by student B? (*b*) When student B throws his football toward student A at 15.0 m/s, it emits a frequency of 670 Hz. What frequency is heard by student A? (*c*) If the two footballs are thrown simultaneously, what is the beat frequency? •••

17.8 Shock waves

65. A jet ski travelling across the surface of water creates a shockwave that makes an angle of 39° with the horizontal. If the waves it creates in the water have a speed of 9.5 m/s, how fast is the jet ski moving? •
66. A military jet is flying at 420 m/s horizontally. Take the speed of sound to be 344 m/s. What is the apex half angle of the conical shock wave it created? •
67. A speedboat traveling at 73.0 km/h causes a shock wave that makes an angle of 14.3° with the path of the boat. At what speed are the waves moving through the water? •
68. Two jets flying at the same altitude pass over your head simultaneously, one traveling at Mach 1.5 and the other at Mach 2.5. Which plane's sonic boom do you hear first? ••
69. While attending an air show, you observe a fighter jet traveling horizontally at a speed of Mach 1.30. At the instant you hear the sonic boom, what is the angle between your line of sight to the jet and the horizontal? ••
70. Many years ago, test pilots accelerating through Mach 1 reported that the ride was pretty rough just before Mach 1 but smoothed out suddenly as Mach 1 was breached. How can you account for this? ••
71. Concerned about being disturbed by sonic booms, residents near an Air Force base located along a seacoast submit a petition asking that all jet aircraft from the base head out to sea before attaining Mach 1 and proceed inland only after they have reached supersonic speed. What do you think of the physics reasoning behind this idea? ••
72. A woman standing on the ground observes a jet directly overhead flying at an altitude of 20,000 m. If the jet has a speed of Mach 2 and its shock wave makes an angle of 30° with the horizontal, how long will it be until she hears the sonic boom? ••
73. At a pistol range, a microphone is placed between a shooter and the target so that when a bullet is fired we can hear its sonic boom. A bullet passes by the microphone at a speed of Mach 2.1. The microphone picks up the sonic boom 0.0021 s after the bullet passes above the microphone. (*a*) What is the angle of the shock wave? (*b*) What is the vertical distance between bullet and microphone? ••

74. A jet flying directly over you at an altitude of 3000 m produces a shock wave. If the angle of the shock wave is 42°, (*a*) how long will it be until the sonic boom reaches you, and (*b*) how far does the jet travel during this time interval? ●●
75. At White Sands, New Mexico, a car powered by a rocket engine passes you while you are standing still. If you hear the sonic boom 0.045 s after the car passes and the angle of the shock wave is 37.0°, what is your perpendicular distance from the path of the car? ●●●

Additional Problems

76. What does an observer hear on the ground before and after a sonic boom caused by a jet passing overhead? ●
77. Is a second sonic boom generated when a plane reaches a speed of Mach 2? ●
78. Two test pilots are flying in a jet trainer going faster than Mach 1. Can the pilot in the front of the cockpit hear the pilot in back? ●●
79. Two guitar strings, 1 and 2, plucked simultaneously produce standing waves described by

$$D_1 = A\sin(ax)\cos(bt)$$
$$D_2 = A\sin(qx)\cos(dt),$$

with $a = 14.5\ \text{m}^{-1}$, $b = 2512\ \text{s}^{-1}$, $q = 19.3\ \text{m}^{-1}$, and $d = 2575\ \text{s}^{-1}$. (*a*) What is the frequency of each wave? (*b*) What is the beat frequency of the combined wave? ●●
80. A race car traveling at 350 km/h passes a spectator in the stands. If the spectator has an air horn that emits a sound of frequency 400 Hz, what is the range of frequencies of the sound heard by the driver of the car? ●●
81. Which type of conic section (circle, ellipse, parabola, or hyperbola) is formed by the nodal-line pattern produced by two coherent sources of two-dimensional waves? (Hint: For a given nodal line, the difference in distance between any point on it and the two sources is the same for all points.) ●●
82. Why is the high-frequency "tweeter" on a stereo speaker usually much smaller than the low-frequency "woofer"? ●●
83. You are standing beneath a Ferris wheel when two children riding in cars on opposite sides of the wheel both scream at 600 Hz. Your well-trained ears notice a beat frequency that is maximum when one child is moving directly toward you and the other is moving directly away from you. The Ferris wheel takes 24.0 s to complete one rotation, and its diameter is 27.0 m. You begin to wonder whether a Doppler shift could produce observable beats under these circumstances. ●●● CR
84. You have tickets for an outdoor rock concert, fourth row from the stage. However, at that distance the intensity level is 100 dB, too loud to be enjoyable. You decide to move to a row where the intensity level is a more moderate 80 dB. Fortunately, turning around and counting the right number of rows back (assuming that the distance between rows is uniform), you see that the seats you want are available. ●●● CR
85. You want to build a portable device that can detect pirate radio stations broadcasting at about 50 W of power a few miles from your beachfront home. You have a 0.60-m diameter bowl you might use as an antenna dish, but you are worried that your amplifier input will require a signal of at least 0.10 μW. ●●● CR
86. Your car is third in line at a railroad crossing as a train moving at a constant speed approaches, with its whistle blowing constantly. To keep yourself from going crazy with the noise, you begin thinking about the relative difference between the frequency you hear and the frequency the engineer hears. Surely this must depend on the angle that your line of sight to the engine of the train makes with the perpendicular to the track. ●●● CR



18

Fluids

CONCEPTS

18.1 Forces in a fluid
18.2 Buoyancy
18.3 Fluid flow
18.4 Surface effects

QUANTITATIVE TOOLS

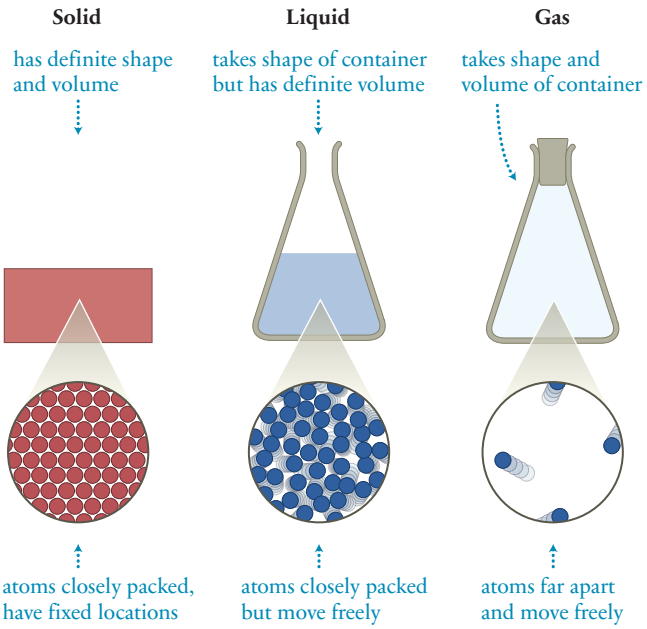
18.5 Pressure and gravity
18.6 Working with pressure
18.7 Bernoulli's equation
18.8 Viscosity and surface tension

In earlier chapters we dealt with the motion and behavior of solid objects mostly. In the next few chapters we extend our description to include two other states of matter: liquids and gases (Figure 18.1). The internal forces exerted on the atoms in a solid keep each atom at a particular location in the solid. Because the atoms can move only slightly relative to one another, a solid material free of external forces has a definite size and shape. In a liquid, the forces exerted on the particles* are not strong enough to keep them from moving relative to one another. Consequently a liquid sample has a definite volume but adjusts to the shape of its container. In a gas, the particles are far apart. They move about randomly and freely, and so a sample of gas has neither a definite volume nor a definite shape.

Liquids and gases are collectively called **fluids** because they flow.

Because fluids flow and change shape, we cannot describe their behavior in the same way we described the behavior of solids. When you look at the flow of water in a river, for example, it makes no sense to speak of "the velocity of the water" because the water may move fast in one spot and more slowly in another, or it may be flowing in one direction at one location and in another direction a bit farther downstream. Unlike a solid object, in other words, water in a river cannot be characterized by a simple set of kinematic quantities—the motion of its center of mass and rotation about that center. However, you can specify its velocity at a specific location.

Figure 18.1 States of matter.



*Both liquids and gases can be made up of either atoms or molecules. From now on, we'll use the generic term *particle* to refer to either atoms or molecules in a liquid or gas.

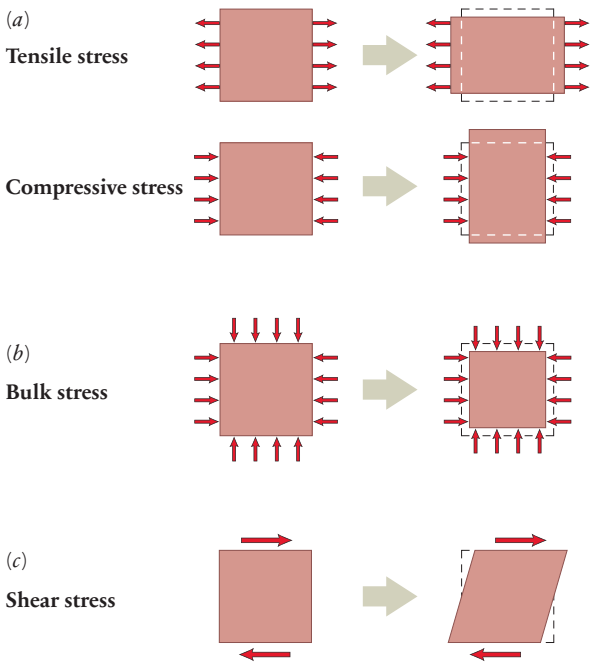
In general, therefore, describing the motion of fluids requires an approach in which we associate quantities such as velocity not with a body of fluid as a whole but with a position within the body of fluid.

18.1 Forces in a fluid

The result of exerting forces on fluids is very different from the result when forces are exerted on solids. To better understand this difference, let us begin by examining how a solid object at rest responds to a set of forces whose vector sum is zero. Because the vector sum is zero, the object remains at rest, but because no solid material is perfectly rigid, the object deforms when opposing forces are exerted on it. Such opposing forces are said to cause a *stress*. A solid object that returns to its original shape when the stress is removed is said to be *elastic* (see Section 8.6). Beyond the elastic limit, the object deforms permanently, and still greater forces may cause the object to fracture. Within the elastic range, the amount of deformation caused by a given stress depends on the properties of the material and typically obeys Hooke's law (see Section 8.9) for small deformations; that is, the deformation is linearly proportional to the magnitude of the applied forces.

Because the atoms in a solid are at fixed positions, a solid object can withstand forces applied to its surface in a number of ways. A set of opposing forces along the same line of action (Figure 18.2a) causes either *tensile stress*, which stretches the object, or *compressive stress*, which compresses it. Under such stress, the object's dimensions perpendicular to the line of action of the forces also change to compensate for the stretching or

Figure 18.2 Different types of stresses caused by equal forces applied in opposite directions to solid objects.



CONCEPTS

compressing along the line of action. In spite of this compensating change, however, precise measurements show that the object's volume increases by a small amount when the object is stretched and decreases by a small amount when the object is compressed.

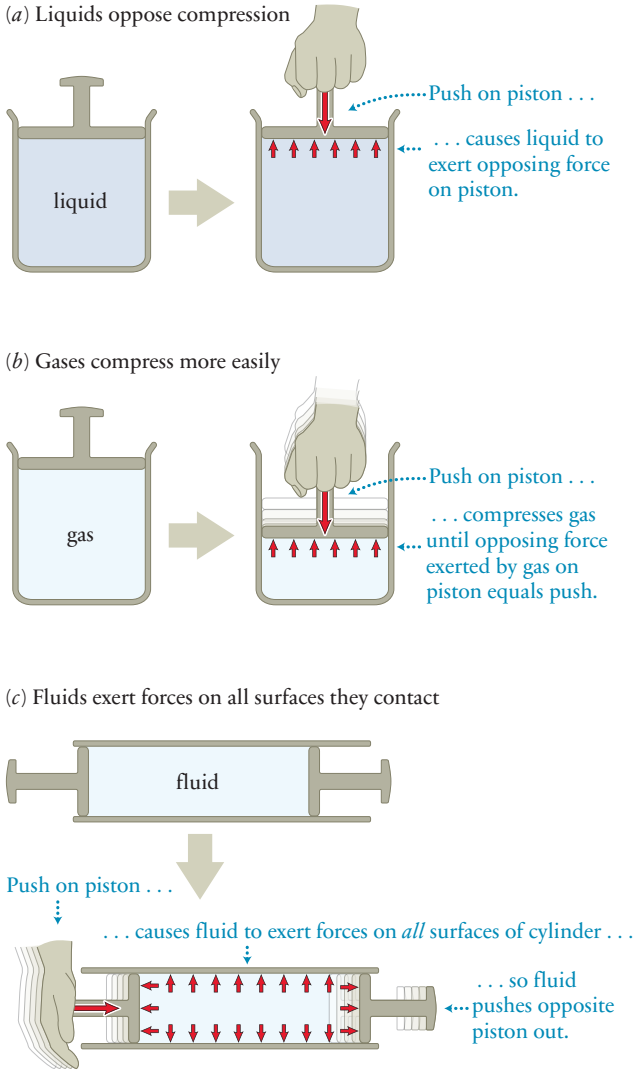
When opposing forces are exerted perpendicular to all surfaces of a solid object (Figure 18.2*b*), the atoms are pushed closer together in all directions, resulting in what is called *bulk stress*. Such a stress causes the volume of the object to decrease.

A third type of stress, *shear stress*, occurs when opposing forces are exerted tangent to the surfaces of a solid object (Figure 18.2*c*). For small forces, the object's shape changes—turning a rectangular face into a parallelogram, say—but the volume of the object does not change.

Let us now examine how a fluid responds to these same three types of stresses: tensile/compressive, bulk, and shear. If you put your hands under water parallel to each other and some distance apart and then move them toward each other, you feel resistance. Your moving hands exert a force on the water and, because forces always come in interaction pairs, the water, in turn, exerts a force on your hands. Once your hands stop moving, however, this force exerted by the water on them disappears. In addition, when you remove your hands, the water molecules don't flow back to where they were before you moved your hands. In other words, water is not elastic the way some solid materials are. If you compress an elastic solid—a foam rubber ball, say—the way you "compress" the water by bringing your hands together, you must keep exerting the force if you want to keep the ball compressed. Once you remove your hands, the ball returns to its original shape. This is not true for most liquids (water, rubbing alcohol, and gasoline, for instance) or for all gases under ordinary circumstances; these fluids are not elastic. You can easily verify that they also cannot support static tensile or shear stresses. Such fluids are said to be *nonviscous*. Liquids such as jelly or tar, which do not behave the way water does in response to a shear stress, can support a shear stress and are said to be *viscous*.

Let us next look at how fluids behave when subjected to bulk stress. If you exert forces on a fluid so as to cause a bulk stress, as in Figure 18.2*b*, the fluid resists being compressed. Imagine putting water in a cylinder fitted with a piston and pushing down on the piston (**Figure 18.3*a***). Because the water has nowhere to go and because the water molecules are already as close together as they can get, a very large force is required to change the volume of the water, and so you feel a rapidly growing resistive force exerted by the water on the piston. If, instead of water, you put a gas in the cylinder, the gas compresses easily as you push on the piston because gas particles are far apart and can easily be pushed closer together. As the gas is compressed, however, it exerts on the piston a force of increasing

Figure 18.3 All fluids resist bulk stresses, but liquids don't compress much, whereas gases are easily compressed.



magnitude. Once that force is equal in magnitude to the force you exert on the piston, the system reaches mechanical equilibrium (Figure 18.3*b*). The greater the force you exert, the more the gas is compressed.

So, a fluid can sustain forces that are perpendicular to its surfaces, provided the forces are exerted simultaneously on *all* its surfaces. If you pushed the left piston in Figure 18.3*c* inward without also pushing the right piston inward, for example, you would simply push the water and the right piston out of the tube. Likewise, if you replaced a section of either the top or bottom wall of the cylinder in Figure 18.3*c* with a piston and again pushed on the left piston, you would also need to push both the right piston and the newly added piston inward in order to keep the water contained. From these observations we conclude that, in mechanical equilibrium, a fluid in a container exerts a (normal) force that is distributed along *all* surfaces of the container that are in contact with the fluid, and those

CONCEPTS

surfaces exert on the fluid a force that is of equal magnitude but in the opposite direction.

The main difference between gases and liquids is that gases tend to compress easily, often greatly changing the volume and therefore the mass density (mass per unit volume, see Section 1.6), whereas the volume and mass density of a liquid do not change easily. For simplicity, from now on we shall assume that the liquids we deal with are *incompressible*; that is, their mass density and volume are fixed.

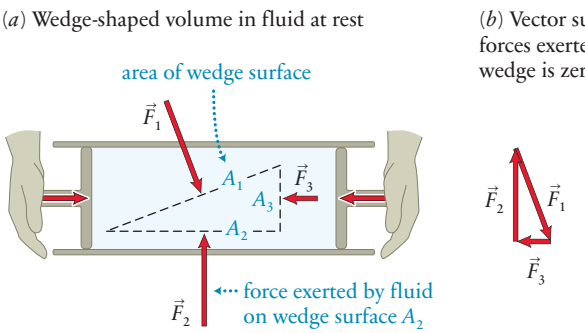
When a gas is subjected to bulk stress, its mass density and volume change easily. When a liquid is subjected to bulk stress, its mass density and volume can be considered unchanged.

18.1 In Figure 18.3*c*, suppose you push the left piston rightward with a force of magnitude 10 N. (For now, ignore normal forces.) (*a*) To keep the water in the cylinder at rest, what force must you exert on the (identical) right piston? (*b*) How does the force $\vec{F}^{c}_{wr}$ exerted by the water on the right piston compare to the force $\vec{F}^{c}_{pl}$ exerted by you on the left piston?

The two pistons in Checkpoint 18.1 are identical. Suppose we divide the right piston in Figure 18.3*c* into two pistons, each having half the surface area of the original piston. As you might expect, to keep the water in the cylinder at rest while pushing the left piston rightward with a force of magnitude 10 N, you would need to exert on each half-piston a force of magnitude 5 N, which suggests a linear proportionality between surface area and force.

To examine this relationship in more detail, consider a fluid at mechanical equilibrium in a cylinder (**Figure 18.4**). Forces are exerted on the two pistons so that the fluid is subject to a compressive stress. Now imagine a wedge-shaped volume of the fluid. Because liquids cannot sustain any shear stresses (resulting from forces tangent to the wedge surfaces) exerted on the wedge, the forces exerted by the surrounding fluid on the wedge must be perpendicular to the three wedge surfaces. Because the wedge is at rest, the vector sum of these forces must be zero. Let the three wedge surfaces have areas A_1, A_2, and A_3. Each of these areas is proportional to the length of the corresponding side of the dashed triangle in Figure 18.4*a*, and each is subject to a normal force exerted by the surrounding fluid. Let us denote these forces by $\vec{F}_1$, $\vec{F}_2$, and $\vec{F}_3$. Because these vectors add up to zero, we can arrange them in a triangle as shown in Figure 18.4*b*. Because the vectors are perpendicular to the sides of the dashed triangle in Figure 18.4*a*, the triangle in Figure 18.4*b* is similar to the dashed triangle, and so the ratio of the force magnitudes $F_1{:}F_2{:}F_3$ is the same as the ratio of the lengths of the sides of the dashed triangle. Because the dimension of the wedge in the direction perpendicular to the plane of the drawing is fixed, the area A of each side of the wedge is proportional to its length. Therefore $F_1{:}F_2{:}F_3 = A_1{:}A_2{:}A_3$, or $F_1/A_1 = F_2/A_2 = F_3/A_3$. In other words, the ratio of the magnitude of the force exerted by the fluid on a surface and the area of that surface is constant. This ratio is called **pressure,** denoted by the symbol P. Pressure is a scalar and is taken to be positive when a fluid is under compression.

Figure 18.4 Forces exerted on the sides of a wedge-shaped volume of fluid.



It is important to distinguish pressure from force. The two are related because forces exerted on the fluid cause the pressure in a fluid. Force, however, is a vector, while pressure is a scalar. In addition, forces of the same magnitude can cause very different pressures depending on the surface area over which these forces are exerted. Imagine, for example, exerting a force of 10 N on your skin first with the tip of a pin and then with the cover of a book. Even though the force is the same in both cases, the effect is very different. At the tip of the pin, the pressure caused by the 10-N force is very great because the surface area of the pin is small. The pressure caused by the pin is likely to cause pain, and the pin might penetrate into your skin. The book, however, spreads that same 10-N force out over a much greater surface area, and so the resulting pressure is much smaller.

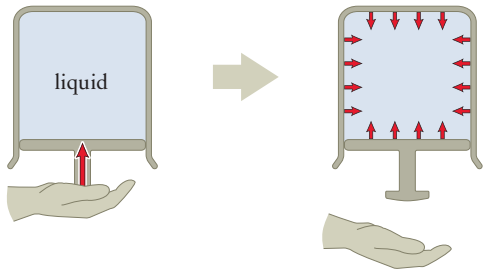
Another important point to keep in mind is that because forces always come in interaction pairs, the force exerted by a container wall on a fluid (which causes the pressure in the fluid) and the force exerted by the fluid on the container wall form an interaction pair. The pressure in a fluid therefore causes the fluid to exert forces on anything that is in contact with it. Because of the mobility of the particles in a fluid, pressure is transmitted in all directions to all parts of the fluid. However, because fluids cannot sustain shear stresses, the forces exerted by the fluid on any surfaces that contain the fluid are always exerted perpendicular to the surface.

In gases, pressure can be only positive. Liquids, however, can sustain negative pressures. For example, if you hold the container of Figure 18.3*a* upside down, as shown in **Figure 18.5*a*** on the next page, and pull down on the piston, you feel a large resistive force, which tells you that the liquid pulls the piston inward. The water molecules exert an attractive force on one

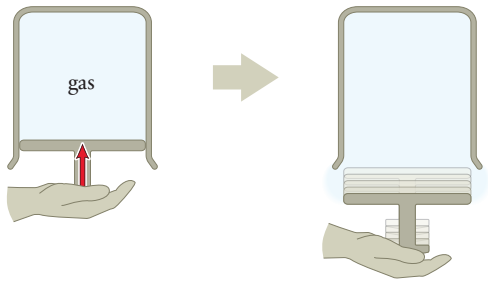
CONCEPTS

Figure 18.5 Liquids can sustain negative pressures, whereas pressures in a gas can be positive only.

(*a*) Because liquid particles exert attractive forces on each other and on the container walls, liquid can hold piston up against force of gravity



(*b*) Gases expand to fill any volume and therefore cannot sustain negative pressures



another and on the surfaces of the container, pulling them inward as you pull the piston down. So, just as a liquid under compression causes an outward push that we associate with a positive pressure, we find that a "stretched" liquid causes an inward pull that we associate with a negative pressure.

(You can easily try this yourself. Fill a glass to the top with water, place an index card on top of the glass, making sure no air is left between the water surface and the card, and then turn the glass upside down while holding the card in place with the palm of your hand. If you slowly remove your hand from the card, the water remains in the glass.)

Gas particles are so far apart that they interact only when they collide, and so a gas always expands to fill whatever container it is placed in. A gas can therefore not sustain negative pressure, and a piston just falls out of a gas container that is held upside down (Figure 18.5*b*).

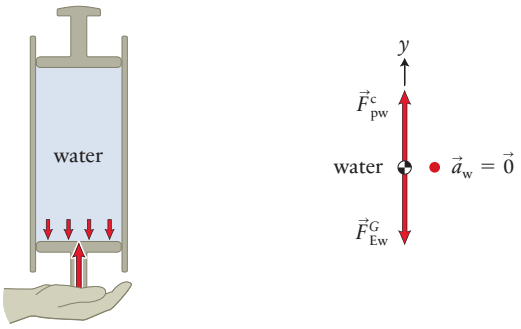
18.2 Consider a plastic cup full of water. (*a*) Is the water exerting a force on the bottom of the cup? (*b*) On the sides of the cup? (*c*) Is the water pressure negative, zero, or positive?

The answers to Checkpoint 18.2 may have surprised you. What causes the pressure in the water if the top surface of the water is open to the environment and no piston is pushing down on that top surface? The cause

Figure 18.6 Gravity exerts a downward force on a volume of fluid.

(*a*) You must exert upward force on bottom piston to keep water from falling out

(*b*) Free-body diagram for water



is gravity. Imagine the column of water in the glass as being a stack of thin "layers" of water. Gravity exerts a force on each layer, which means that each layer pushes down on the layer underneath it, just as a book presses down on the surface of a table because of gravity. At the top of the water column, the atmosphere presses down on the surface, our imaginary top layer.

To examine the effect of gravity on the pressure in a fluid, consider again a water-filled tube fitted with identical pistons on either end, but now we orient the tube vertically (**Figure 18.6*a***). Even with no force exerted on the top piston, you have to exert an upward force on the bottom piston to prevent the water from falling out of the tube. As the free-body diagram in Figure 18.6*b* shows, when you are not pressing down on the top piston, the water is subject to two forces: a downward force of gravity and an upward force of equal magnitude exerted by the bottom piston.* So the pressure at the top of the water is zero (no force exerted on the surface), whereas the pressure at the bottom of the tube is nonzero.

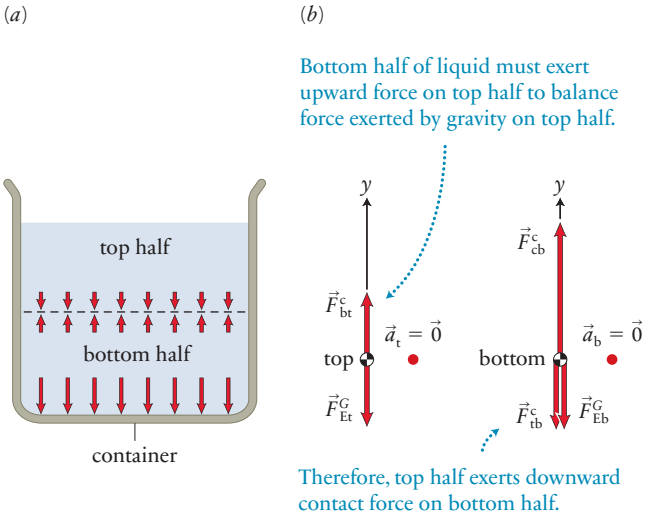
Now imagine dividing a liquid in a container into two equal parts, as shown in **Figure 18.7*a***. Ignoring any force exerted at the top surface (we'll return to this issue in a bit), we see from the free-body diagrams in Figure 18.7*b* that the top half must be supported against the pull of gravity by the bottom half. The bottom of the container, in turn, must support both halves. Consequently the pressure at the bottom of the container is greater than the pressure at the center.

18.3 How does the pressure in the liquid vary with height above the bottom of the container in Figure 18.7?

*As we shall see later, the atmosphere exerts a force on each piston, but the combined effect of these two forces is negligible. For simplicity, we also ignore the mass of the pistons, which means there is no gravitational force exerted on them and so the top piston exerts no force on the water.

CONCEPTS

Figure 18.7 Pressure increases with distance below the surface because liquid layers at the top press down on lower layers of liquid.



The conclusion you reached in Checkpoint 18.3 holds for any liquid that is at rest in a container of any shape:

The pressure in a liquid at rest in a container decreases linearly with height, regardless of the shape of the container.

In a gas, the pressure also decreases with height, but the effect is much less pronounced because gases are much less dense than liquids. Also, because gases are compressible, gas density varies with height and so the decrease in pressure with height is not linear. Air, for example, is densest at sea level, where it is compressed by all the air above. At the cruising altitude of most airplanes (10 km), the mass density of the air is one-third of its value at sea level, and the atmospheric pressure is just one-quarter of its value at sea level.

At sea level, you are at the bottom of an "ocean" of air. The mass of a column of air that has a cross section the size of a fingernail is approximately 1 kg. In other words, at sea level the atmosphere exerts on each fingernail-sized area of your body a force that is equivalent to the force of gravity on a 1-kg object. If we were to add up the magnitudes of the forces exerted on the entire surface area of your body, you would obtain a force equal to the force of gravity exerted on 2×10^4 kg! The force with which the atmosphere pushes down on the surface of your orange juice at breakfast is more than four times the force with which this book presses down on your lap.

Given that the average atmospheric pressure at sea level is 1×10^5 N/m^2, it is clear that we cannot ignore the pressure at the top of the liquid in Figure 18.7. As you saw in Checkpoint 18.1, any force exerted at one surface of a liquid is transmitted to any other part of the liquid. This means that if we change the pressure in a container of liquid at any location, the change in pressure is the same throughout the liquid. This is called **Pascal's principle:**

A pressure change applied to an enclosed liquid is transmitted undiminished to every part of the liquid and to the walls of the container in contact with the liquid.

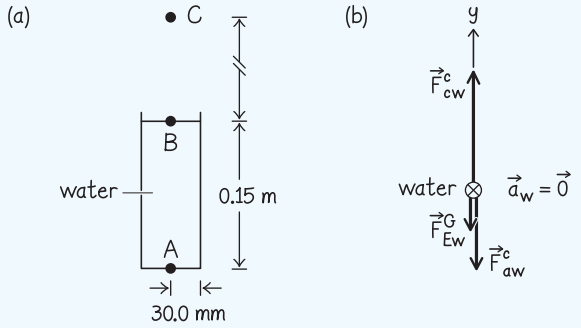
Consequently, if atmospheric pressure at the top of a column of water in a container open to the atmosphere is 1×10^5 N/m^2, the pressure in the water near that surface must also be 1×10^5 N/m^2. Below the surface, the pressure in the water increases with greater depth as more water pushes down due to gravity. The next example gives you some feel for pressure variation with height in both water and air.

EXAMPLE 18.1 Water and air pressure

Consider a cylindrical container that has a radius of 30.0 mm and is filled with water to a height of 0.150 m. If the container is at sea level, what is the pressure (*a*) at the water surface and (*b*) at the bottom of the container? (*c*) How far above the water surface is the decrease in atmospheric pressure the same as the pressure decrease between the bottom and top of the water? The mass densities of air and water are 1.20 kg/m^3 and 1.00×10^3 kg/m^3, respectively.

1 GETTING STARTED I begin by making a sketch of the situation (**Figure 18.8*a***). The problem deals with the pressure at three locations: at the bottom of the container (point A), at the water surface (B), and at some location above the surface (C). Because 0.150 m is a negligible distance in the context of this problem, the pressure at B is essentially equal to the atmospheric pressure at sea level. At A the pressure should be greater than at B because of the force of gravity exerted on the water. At C the pressure should be smaller than at B because the volume of air above C pressing down on C is smaller than the volume of air above B pressing down on B.

Figure 18.8



2 DEVISE PLAN The pressure at A is equal to the magnitude of the force $\vec{F}^c_{cw}$ exerted by the container bottom on the water divided by the area of the container bottom. To determine the magnitude of this force, I make a free-body diagram for the water (Figure 18.8*b*). The forces exerted

(Continued)

CONCEPTS

on the water are the force of gravity $\vec{F}^G_{Ew}$, the downward force $\vec{F}^c_{aw}$ exerted by the atmosphere, and the upward force exerted by the container bottom, which I need to determine. Because the water is at rest, the vector sum of these forces must be zero, and so the magnitude of the force exerted by the container bottom on the water is $F^c_{cw} = F^G_{Ew} + F^c_{aw}$. The force $\vec{F}^c_{aw}$ is determined by atmospheric pressure. To calculate $\vec{F}^G_{Ew}$ I need to know the mass of the water, which I can obtain by multiplying the volume of the water by the mass density of water.

Because the difference in pressure between A and B is the same as that between B and C, I know that the mass of the air column between B and C must be equal to the mass of the water column between A and B. The ratio of the distances between these locations must therefore be equal to the ratio of the volumes, which is equal to the inverse of the ratio of the mass densities (assuming that the density of air is approximately constant over that column of air).

3 EXECUTE PLAN (*a*) The pressure at B is essentially equal to atmospheric pressure at sea level, which I can look up: $P_B = 1.01325 \times 10^5\ \text{N/m}^2$. ✔

(*b*) Dividing both sides of $F^c_{cw} = F^G_{Ew} + F^c_{aw}$ by the bottom surface area A, I get $P_A = F^G_{Ew}/A + P_{atm}$ because $P_A = F^c_{cw}/A$ and $P_{atm} = P_B = F^c_{aw}/A$. The volume of the water is $V = Ah$, and so the mass of the water is this volume times the mass density of water: $m_{water} = Ah\rho$. The magnitude of the force of gravity exerted on the water is $F^G_{Ew} = m_{water}g = Ah\rho g$, and this force is exerted over the bottom surface area A of the container. The pressure at point A is thus

$$P_A = \frac{F^G_{Ew}}{A} + P_{atm} = \frac{Ah\rho g}{A} + P_{atm} = h\rho g + P_{atm}$$
$$= (0.150\ \text{m})(1.0 \times 10^3\ \text{kg/m}^3)(9.8\ \text{m/s}^2) + 1.01325 \times 10^5\ \text{N/m}^2 = 1.03 \times 10^5\ \text{N/m}^2. \checkmark$$

(*c*) The ratio of the distances is $\overline{AB}/\overline{BC} = \rho_{air}/\rho_{water} = (1.20\ \text{kg/m}^3)/(1.00 \times 10^3\ \text{kg/m}^3) = 1.20 \times 10^{-3}$. I am given $\overline{AB} = 0.150\ \text{m}$, so $\overline{BC} = (0.150\ \text{m})/(1.20 \times 10^{-3}) = 1.25 \times 10^2\ \text{m}$. ✔

4 EVALUATE RESULT The answer I obtained in part *b* shows that the pressure difference between the top and bottom of the container is very small. That makes sense—I know from experience that my ears, which are very sensitive to pressure, don't experience any pressure difference when I swim a mere 0.15 m under the surface of water.

The answer to part *c* tells me that the distance above sea level I would have to climb in order to obtain the same pressure difference is about three orders of magnitude greater than the distance from B to A (0.15 m versus 120 m). Again, I know from experience that I don't experience any detectable pressure difference when going up 120 m in air, so my answer is not unreasonable.

18.4 How far below the surface of a deep body of water is the pressure twice what it is at the surface?

18.2 Buoyancy

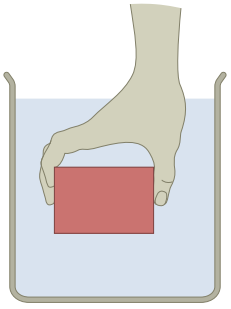
The decrease in pressure with height is responsible for the **buoyant force,** an upward force experienced by objects when they are either fully or partially submerged in a fluid. The tendency of certain objects to float in air or water, called *buoyancy*, is due to this force.

If you hold a brick under water with its top surface parallel to the surface of the water, as illustrated in Figure 18.9*a*, the pressure in the water at the bottom of the brick is greater than the pressure at the top, so the upward force exerted by the water on the bottom surface of the brick is greater than the downward force the water exerts on the top surface of the brick. The free-body diagram for the brick in Figure 18.9*b* shows the forces exerted on the brick: a downward gravitational force, an upward force exerted by your hand, an upward force exerted by the water on the bottom surface, and a downward force exerted by the water on the top surface. Because the pressure is greater at the bottom of the brick than at the top, the two forces exerted by the water on the brick add up to an upward buoyant force

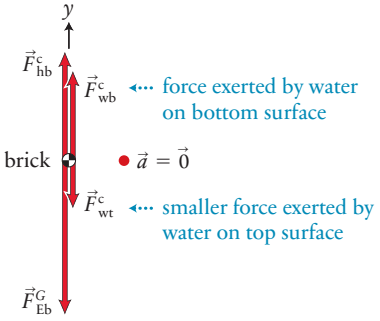
CONCEPTS

Figure 18.9 Cause of the buoyant force experienced by objects that are partly or fully submerged in a fluid.

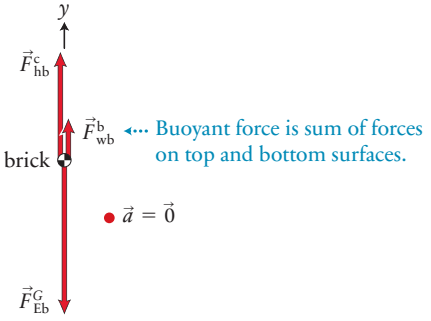
(*a*) Brick held under water



(*b*) Free-body diagram that separates force exerted by water on brick's top and bottom surfaces



(*c*) Free-body diagram showing buoyant force exerted by water on brick



(Figure 18.9*c*). Buoyant forces are denoted by a superscript b. (We can ignore the sideways forces exerted by the water on the brick because they cancel out.)

18.5 (*a*) What happens to the pressure difference between the top and bottom of the brick in Figure 18.9*a* when the brick is held deeper in the water? (*b*) What happens to the pressure difference when the pressure at the surface of the water is increased? (*c*) What is the effect of the pressure in the water on the vertical sides of the brick? (*d*) Consider the same brick held in the air. Does the air exert a buoyant force on the brick? If so, in which direction is this force?

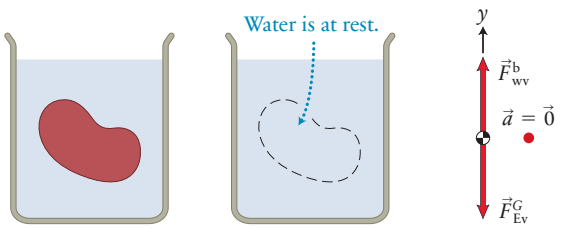
How large is the buoyant force that is exerted on an object submerged either fully or partially in a fluid? For an object that has a rectangular cross section, such as the brick in Figure 18.9*a*, we can determine the pressures at the top and bottom of the object, multiply the pressures by the surface area to get the two force magnitudes, and then take the difference of these magnitudes to obtain the buoyant force. However, many objects do not have a rectangular cross section. Consider, for example, the fully submerged object in **Figure 18.10*a***. Even if we know the pressure at every location on the surface of the object, calculating the buoyant force exerted by the water on the object is a complicated matter. We know, however, that if we remove the object from the water, water fills the space previously occupied by the object (Figure 18.10*b*). This volume of water within the boundary defined by the object is said to be *displaced* by the object when the object is fully submerged. Figure 18.10*c* shows the free-body diagram for this displaced volume of water. Because all the water is at rest, we know that the vector sum of the forces exerted on the volume of water within the boundary must be zero. Two forces are exerted on this volume of water: a downward force of gravity exerted by Earth and a buoyant force exerted by the surrounding water. Given that the vector sum of these two forces is zero, we know that the buoyant force exerted on the water within the boundary must be upward and equal in magnitude to the magnitude of the force of gravity exerted by Earth on the water within the boundary.

Because the buoyant force exerted by the surrounding water on the boundary in Figure 18.10*a* cannot depend on the nature of the material inside the boundary, we have a way of determining the buoyant force exerted on an object of arbitrary shape:

> **An object submerged either fully or partially in a fluid experiences an upward buoyant force that is equal in magnitude to the magnitude of the force of gravity exerted on the fluid displaced by the object. The volume of the displaced fluid is equal to the volume of the submerged portion of the object.**

This statement is called **Archimedes' principle.** Note that there is no reference to the shape or composition

Figure 18.10 The force exerted by gravity on the volume of liquid *displaced* by a submerged object is equal in magnitude to the buoyant force exerted by the liquid on the object.



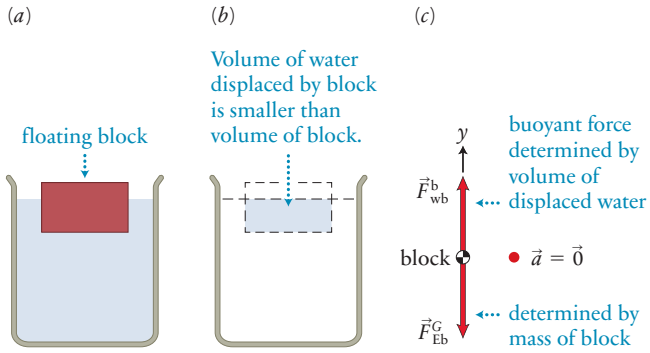
of the submerged object: The only factor that comes into play is the volume of the displaced fluid.

18.6 If the buoyant force exerted on an object is always equal in magnitude to the force of gravity exerted on the fluid displaced by the object, why does a brick placed in a barrel of water sink?

Now that we know how to determine the buoyant force exerted on an object, we can understand why some objects float and others sink. As Checkpoint 18.6 shows, a brick sinks because the force of gravity exerted by Earth on it is greater than the force of gravity exerted by Earth on an equal volume of water (equal in magnitude to the buoyant force). Because the vector sum of these two forces is directed downward, the brick sinks unless it is held up.

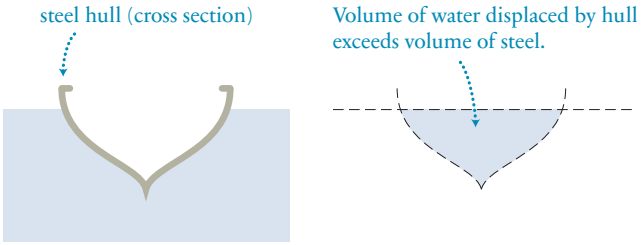
When the buoyant force is greater in magnitude than the force of gravity exerted on an object fully submerged in water, the object accelerates toward the surface of the water when you release it. A block of wood released under water, for example, accelerates to the surface, so the force of gravity exerted by Earth on the block is smaller than the force of gravity exerted on an equal volume of water (**Figure 18.11*a***). Once it floats at the surface, only part of the block is submerged, and so now the volume of displaced water is smaller than the volume of the block (Figure 18.11*b*). Because the block

Figure 18.11 The concept of displaced volume can be used to determine the buoyant force on floating as well as submerged objects.



CONCEPTS

Figure 18.12 Steel ships can float even though steel sinks. The ship's hollow interior means that the volume of water displaced by the ship is larger than the volume of steel. Thus, the upward buoyant force exerted on the ship is greater than the downward gravitational force exerted on the steel.



is at rest at the surface, the force of gravity exerted by Earth on it and the buoyant force exerted by the water must be equal in magnitude, as shown in Figure 18.11*c*.

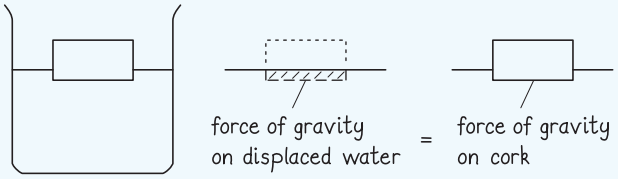
Objects made out of materials that normally sink can be made to float by designing the objects in such a way that the buoyant force is great enough to offset the force of gravity exerted on them. For example, a steel ship floats even though a block of steel sinks. The reason for this is illustrated in **Figure 18.12**: As long as the force of gravity exerted on the ship is smaller than the force of gravity exerted on the volume of water displaced by the ship, the buoyant force is great enough to keep the ship floating. Because the hull of a ship is hollow, the volume of displaced water is much greater than the volume of steel.

EXAMPLE 18.2 Mass densities

When a piece of cork floats in water, three-quarters of its volume is above the water surface. How does the mass density of cork compare with the mass density of water?

❶ **GETTING STARTED** I begin by making a sketch of the situation. Because one-quarter of the cork's volume is under water, I know that the cork displaces a volume of water equal to one-quarter of the volume of the cork. This also tells me that the force of gravity exerted on the whole cork $m_{\text{cork}}g$ is equal to the force of gravity exerted on the displaced water $m_{\text{disp}}g$ (**Figure 18.13**), which means that the mass of the displaced water is equal to the mass of the cork.

Figure 18.13



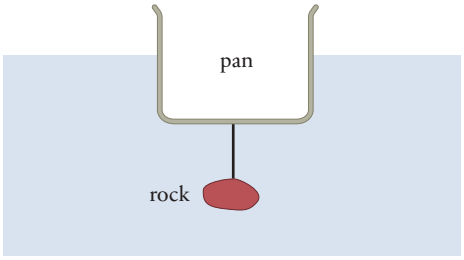
❷ **DEVISE PLAN** The mass density of an object is equal to the amount of mass per unit volume. Because I know how the volumes of equal masses of cork and water compare (Figure 18.13), I can compare the mass densities of cork and water.

❸ **EXECUTE PLAN** The volume of a piece of cork of a certain mass is four times greater than the volume of an equal mass of water, and so the mass density of cork is one-quarter that of water. ✔

❹ **EVALUATE RESULT** A "light" material such as cork has less mass per unit volume and therefore a smaller mass density than water, as I expect.

18.7 A pan with a rock in it is floating in water. Suppose you remove the rock and use a very lightweight string to suspend it from the bottom of the pan (**Figure 18.14**). If the combination still floats, is the volume of water displaced when the rock is suspended greater than, equal to, or smaller than the volume of water displaced when the rock is in the pan?

Figure 18.14 Checkpoint 18.7.



18.3 Fluid flow

We have so far restricted our study to fluids at rest. Now we turn our attention to fluids in motion; such motion is usually referred to as *fluid flow*. When the flow is such that the velocity of the fluid at any given location is constant, such as when water moves through a pipe, the flow is said to be **laminar.** Not all fluid flow is laminar, however. For example, the motion of rising smoke or the motion of water in the wake of a large ship is characterized by chaotic changes. Such flow is said to be **turbulent.**

To better understand fluid flow, we introduce the concept of **streamlines,** which are lines drawn to represent the paths taken by the particles in a fluid in motion. Streamlines can be made visible by injecting smoke into a flow of gas (**Figure 18.15**) or ink into a flow of liquid

Figure 18.15 Streamline flow over a car.



CONCEPTS

Figure 18.16 Streamline flow around an airplane wing.

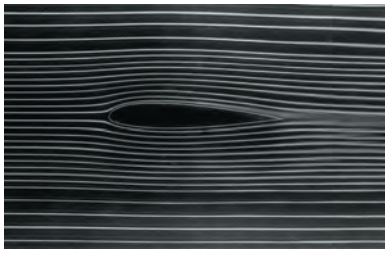
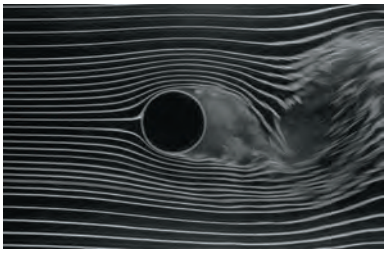


Figure 18.17 Turbulent flow around a cylinder.



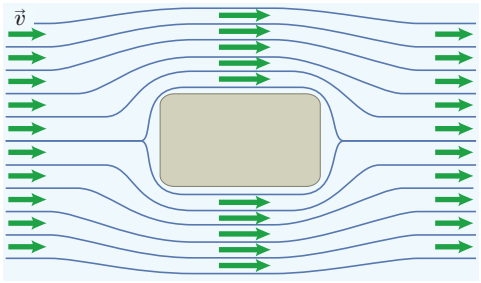
(**Figure 18.16**). When flow is laminar, as in Figures 18.15 and 18.16, streamlines maintain their shape and position. In turbulent flow, streamlines are erratic and often curl into whirling, circularly moving regions of fluid called *vortices* (singular *vortex*, **Figure 18.17**).

Whether a flow of fluid past a stationary object is laminar or turbulent depends on the speed of the flow, the shape of the object, and the fluid's resistance to shear stress (its *viscosity*). For a fluid that has zero viscosity and that has no friction when flowing past surfaces, streamlines just bend around a blunt object (**Figure 18.18*a***). Except where the fluid has to move out of the object's way, the flow speed is the same everywhere. In such situations, the streamline pattern in regions downstream from the object is the same as the streamline pattern upstream of the object. (Although such "totally nonviscous" fluids don't exist, examining these models provides useful insights.)

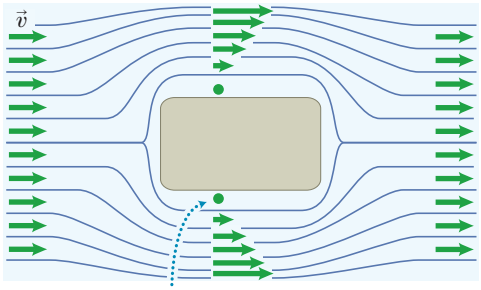
Figure 18.18*b* shows the flow pattern for a fluid that has nonzero viscosity when that fluid is moving at a very low speed. The streamlines have the same general shape as for the nonviscous fluid of Figure 18.18*a*, but the pattern of speeds is very different in the two cases. At the object's surface, the viscous fluid in Figure 18.18*b* is stationary because of friction between the fluid and the surface. Far above and below the object, the flow is unaffected by the presence of the object. Between those two extremes, the flow speed increases from layer to layer. Because the viscosity is nonzero, adjacent layers of fluid exert shear forces on each other, with faster-moving layers trying to speed up slower-moving layers and slower-moving layers

Figure 18.18 Whether the flow of fluid past a stationary object is laminar or turbulent depends on the fluid's speed and viscosity and on the object's shape.

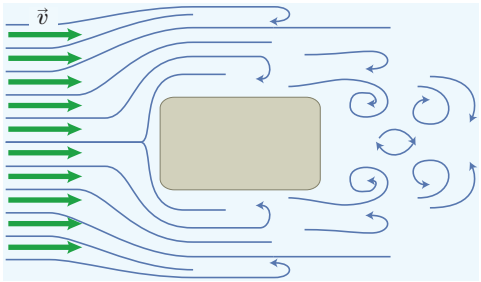
(*a*) Low flow speed, zero viscosity



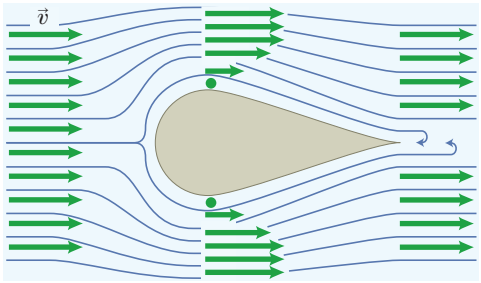
(*b*) Low flow speed, nonzero viscosity



(*c*) High flow speed, turbulence



(*d*) High flow speed, streamlined object



CONCEPTS

slowing down faster-moving ones. The resulting "rubbing" of layers of fluid causes energy dissipation.

As flow speed is increased, the torques caused by the forces that adjacent layers exert on each other become so great that the flow develops vortices and becomes turbulent (Figure 18.18*c*). The turbulent wake alongside and downstream from the object is erratic and causes significant energy dissipation.

18.8 Instead of laminar fluid flow past a stationary object as in Figure 18.18*a*, consider the motion of an object moving at constant velocity through a stationary fluid. Do you expect the flow pattern surrounding the object to be the same as in Figure 18.18*a* or different?

To prevent turbulence and minimize energy dissipation at high flow speeds, it helps to *streamline* objects, which means to adjust their shape so as to maintain streamline flow around them. For example, when an object has a rounded upstream end and a long, tapered downstream tail, as in Figure 18.18*d*, the turbulent wake is very narrow and the flow pattern is similar to the low-speed flows illustrated in Figure 18.18*a* and *b*. At the surface of a streamlined object such as the one in Figure 18.18*d*, the fluid is still stationary, but the flow speed increases rapidly over a thin boundary layer. Outside this boundary layer, the flow resembles that of a nonviscous fluid.

Now that we have some general feeling for fluid flow, let us look at flow through a pipe. In a pipe of fixed diameter, the streamlines are all parallel to the walls of the pipe. For an (ideal) nonviscous fluid flowing through such a pipe, the flow speed is the same everywhere and the pipe offers no resistance to the flow (**Figure 18.19*a***). For any viscous fluid at low flow speed, the flow speed is zero at the pipe walls and highest at the center (Figure 18.19*b*). At high flow speed, the speed is more constant across the diameter of the pipe, dropping off only in a boundary layer (Figure 18.19*c*). The higher the speed, the thinner the boundary layer.

Let us next consider flow through a constriction in a pipe (**Figure 18.20*a***). To simplify matters, we ignore any turbulence or boundary-layer effects. The first thing to note is that the streamlines must crowd together in the narrow part of the pipe. (A streamline cannot have an end because that would mean a fluid particle moving along that streamline would come to a stop.) Therefore the density of streamlines is greater in the narrow part than it is in the wide part.

Figure 18.19 Flow speed patterns through a pipe of fixed diameter.

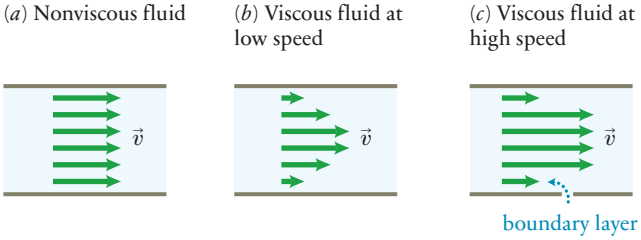
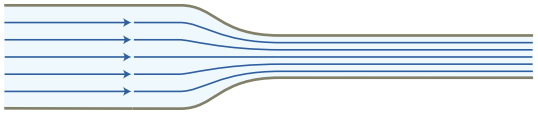
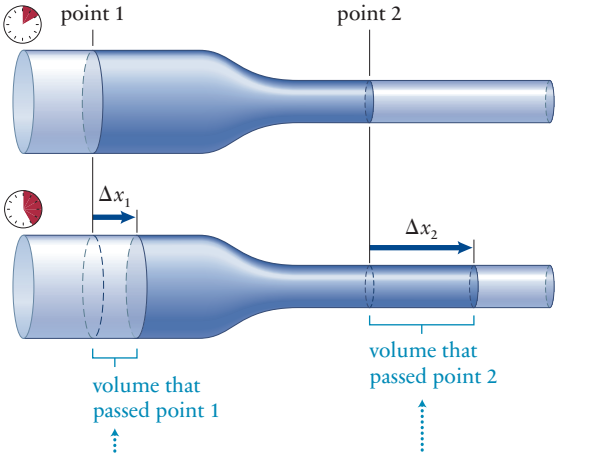


Figure 18.20 What happens when a fluid flows from a wider region of a pipe into a narrower region.

(*a*) When a pipe carrying a fluid narrows, the streamlines get closer together



(*b*) Where the pipe is narrower, the fluid must flow faster



If the flow is laminar, the mass of the fluid passing point 1 in a given time interval in the wide part must be equal to the mass of the fluid passing point 2 in the narrow part in that time interval (Figure 18.20*b*). (If that were not the case, mass would accumulate between those two points and the flow would not be laminar.) Consequently, in a given time interval, the displacement Δx_2 of a fluid particle moving through the narrow part is greater than the displacement Δx_1 of a fluid particle moving through the wide part. This means that the overall fluid flows faster in the narrow part of the pipe where the density of streamlines is greater than in the wide part.

A flowing fluid speeds up when the region through which it flows narrows and slows down when the region through which it flows widens. The density of streamlines reflects the changes in flow speed: When streamlines in a laminar flow get closer together, the flow speed increases; when they get farther apart, the flow speed decreases.

This principle applies to every particle of a moving fluid. Thus each particle must accelerate as it moves

CONCEPTS

along a streamline from the wide part of the pipe to the narrow part. This means that the surrounding fluid must exert a force to the right on the particle while it is in the tapered region of the tube in Figure 18.20. Because the force exerted on fluid particles is a measure of the pressure in the fluid, the pressure is greater in the wide part of the pipe than in the narrow part.

When the flow speed in a laminar flow increases, the pressure in the fluid decreases.

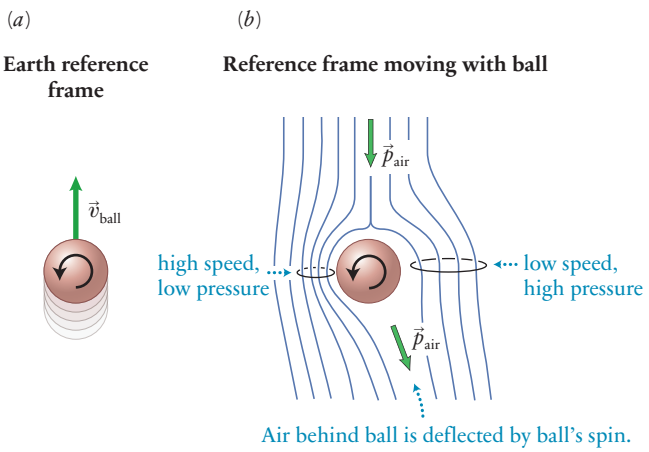
This effect is sometimes called the *Bernoulli effect* after the Swiss physicist who first described the pressure difference quantitatively. The pressure difference occurs along the streamlines. Because there is no flow in the direction perpendicular to a streamline, the pressure must be the same across neighboring streamlines.

There are many examples of the Bernoulli effect. During a storm, fast winds blowing over the roof of a house cause the pressure above the roof to be lower than the pressure in the still air inside the house, creating an upward force on the roof. In extreme circumstances, this force can be great enough to lift the roof off the house. When a large truck speeds past a stationary car, the fast-moving air between the car and the truck has a lower pressure than the still air on the other side of the car, and consequently the car is pulled toward the truck.

Another familiar example is the curved, gravity-defying motion of a rapidly moving, spinning ball. **Figure 18.21***a* shows a top view of a curveball—a baseball thrown with a counterclockwise spin. Because of friction between the surface of the ball and the air, air is dragged around with the spinning ball as the ball moves through the air. Figure 18.21*b* shows the air movement viewed from a reference frame moving along with the ball. The motion of the air caused by the spinning of the ball adds to the flow of air past the ball, causing the air behind the ball to be deflected to the right and the air on the left of the ball to move faster than the air on the right. The deflection of the air means that the direction of the momentum of the air shifts rightward, and that change must be accompanied by an equal-magnitude leftward momentum change for the ball. Alternatively, according to the Bernoulli effect, the pressure in the air to the left of the ball is lower than the pressure to the right, causing a leftward force on the ball. Both of these effects make the ball veer to the left. Likewise a tennis ball that has "top spin" (meaning the top of the ball spins in the direction of the ball's translational motion) curves sharply downward.

The dimples on the surface of golf balls increase the friction between the air and the ball surface to enhance the effect of spin. A golf ball struck with a club is given a rapid backspin because of the slant of the club's face. The backspin forces the air behind the ball downward, causing the ball to experience a lift that carries it farther than it would travel without spin. The lift gained by the backspin more than compensates for the loss of range resulting from air drag.

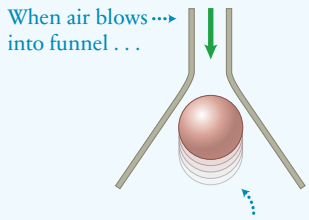
Figure 18.21 Motion of a spinning ball as seen (*a*) from the Earth reference frame and (*b*) from a reference frame moving along with the ball. Because the air is dragged along with the ball surface, the air moves faster to the left of the ball and more slowly to the right of the ball.



EXAMPLE 18.3 Magic pull

When air is blown downward into the narrow opening of a funnel (**Figure 18.22**), a lightweight ball initially positioned below the wide opening is pulled up into the funnel and held in place. Explain how the ball can be pulled upward even though the airflow and gravity are both directed downward.

Figure 18.22 Example 18.3.



❶ **GETTING STARTED** I know that if I were to blow against the ball, it would move away from me. So I reason that to pull the ball upward, there must be exerted on it an upward force that is greater than the combined effect of the downward gravitational and airflow forces. The ball is subject to an upward buoyant force, but I know this force is smaller than that of gravity alone because a ball—even a lightweight one—does not float in air. So there must be an additional force pulling the ball upward. Because this force is absent when no air flows past the ball, that flow must be causing the upward force.

❷ **DEVISE PLAN** To understand the effects of the flowing air, I should sketch the streamlines (I assume laminar flow). The density of the streamlines tells me how the speed of the flow varies, and from this information I can determine how

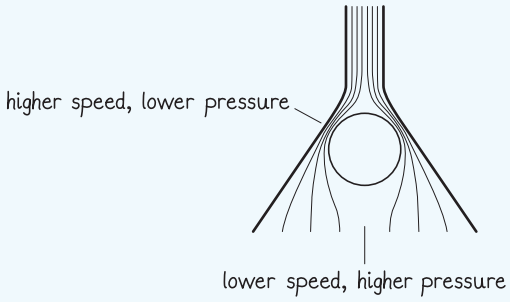
(Continued)

CONCEPTS

the air pressure varies around the ball, which in turn tells me in which direction the air exerts a force on the ball.

❸ **EXECUTE PLAN** **Figure 18.23** shows how the streamlines must bend around the ball. In order to pass around the ball, they must be squeezed together between the ball and the sides of the funnel. At the bottom of the funnel, the diameter of the funnel is very wide and so the streamlines are spaced far apart. This tells me that the flow speed is smaller below the ball than above it and consequently that the pressure is greater below the ball than above it. This pressure difference causes an upward force on the ball. If the pressure difference is great enough, this force can exceed the combined effect of the force of gravity and the downward force of the airflow on the ball. ✔

Figure 18.23



❹ **EVALUATE RESULT** My answer is consistent with the observation given in the problem statement, even though the ball's being pulled *upward* by a downward flow of air seems nothing short of magic! I assumed the flow to be laminar, which may not be entirely correct for an object like a ball (which is not streamlined). However, because air is a gas and because gases have very low viscosity, the assumption that the flow remains laminar is not unreasonable.

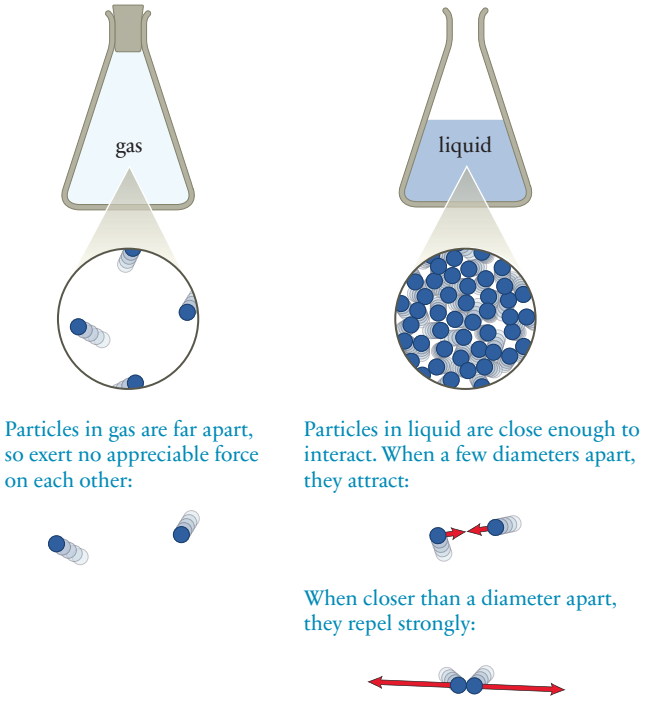
CONCEPTS

18.9 The cloth roof of a convertible car often bulges at high speed, even when the top fits tightly and no wind is getting caught under it. Explain what causes the bulging.

18.4 Surface effects

A distinguishing feature between liquids and gases is that liquids have a well-defined surface but gases do not. If you put a gas in an open container, the gas escapes. A liquid, however, remains in the container it is placed in, and small quantities of liquid placed on a solid surface form drops. In each case, the liquid has a well-defined surface that separates it from the surrounding air. On a small scale—that is, for small quantities of liquid and for small objects—this surface behaves like an elastic skin, pulling small volumes of liquid into drops and supporting small objects such as dust particles or water spiders.

Figure 18.24 Interaction between fluid particles in a gas when the particles are far apart and in a liquid when they are closer together.



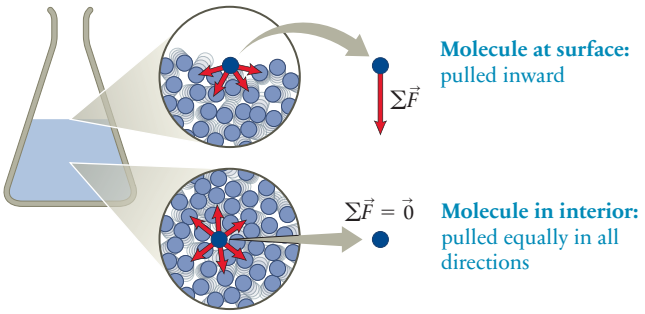
We can understand the behavior of liquid surfaces by considering what happens at the molecular scale. When the particles in a fluid are far apart, they exert no appreciable force on one another (**Figure 18.24**). As they get closer together, however, the interaction between particles has both an attractive and a repulsive component. At a distance of a few particle diameters, particles exert attractive forces, called *cohesive forces*, on one another. When the separation distance becomes less than a particle diameter, however, the particles strongly resist being pushed closer together, resulting in very strong *repulsive forces* between particles. In a gas, the particles move at such great speed and their separation distances are so great that the cohesive forces are negligible. In other words, the gas particles essentially don't interact except when they collide and exert repulsive forces on one another.

Cohesive forces play an important role in liquids, however. The piston in Figure 18.5*a* stays in position because of the cohesive forces in the liquid, for instance. Likewise, if you press pure water with no air dissolved in it between two smooth plates, an enormously large force is required to separate the plates.

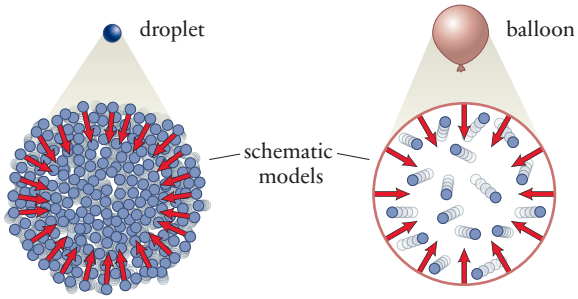
The cohesive forces between particles cause a liquid surface to act like a stretched elastic sheet that tends to minimize the surface area. To see why this happens, compare the forces exerted on a particle in the interior of a liquid with those exerted on a particle at the surface. The interior particle experiences attractive forces

Figure 18.25 The cohesive forces between molecules in a liquid cause a liquid surface to act like a stretched elastic sheet. Consequently, small volumes of liquid tend to adopt a spherical shape.

(*a*) Difference between vector sum of cohesive force on molecules at the surface and in the interior of a liquid



(*b*) Balloon as analogy for surface forces on liquid droplet



Inward pull on surface molecules tends to make droplet spherical.

Stretched balloon skin exerts analogous inward force on gas.

from neighbors in all directions, and the vector sum of these forces is zero. At the surface, however, there are no neighbors above the particle, and so the vector sum of the cohesive forces points into the liquid perpendicular to the surface. The surface particles cannot move into the liquid, however, because as they are pulled inward, they experience a strong repulsive outward force from neighbors.

Because of the inward pull on all surface particles, a small volume of liquid tends to arrange itself in the shape of a sphere (**Figure 18.25**). In a gas there is no attraction between neighboring particles, and therefore a collection of gas particles tends to fly apart, expanding into all available surrounding space. To keep gas particles together, you can put them inside a balloon. The stretched balloon pushes inward on the gas particles, exerting a force on them and keeping them together. The effect of the inward pull on the surface particles in a liquid is much like the force exerted by a balloon on the particles in a gas. The liquid surface thus acts like a membrane that is under tension and that pulls the liquid together in a spherical shape.

Figure 18.26 Astronaut Kevin Ford watches a water bubble float freely between him and the camera aboard the International Space Station. When no large gravitational force is involved, even a large volume of liquid coalesces into a spherical shape.



Just as an inflated balloon is said to be under tension, so do we say that a liquid surface has *surface tension*. In the absence of gravity, surface tension causes even large volumes of liquid to take on a spherical shape (**Figure 18.26**). In the presence of gravity, however, drops of liquid are not perfectly spherical, with larger drops flattening out more than small ones. The reason for the difference in how much large and small drops lose their spherical shape is that only particles on the drop surface are pulled into the spherical shape. The force of gravity, in contrast, pulls on all particles in a volume of liquid, not just those in the surface layer. To state it another way, surface tension is a surface effect but gravity is a volume effect. As the next exercise shows, the relative importance of these two effects changes with drop size.

CONCEPTS

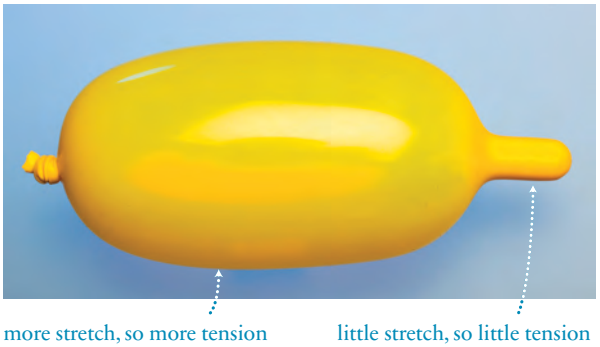
EXERCISE 18.4 Scaling a drop

Suppose surface tension increases in direct proportion to surface area (an oversimplification!). How does the ratio of surface tension to the force of gravity exerted on a drop change as the drop diameter increases tenfold?

SOLUTION The force of gravity is proportional to the mass of each drop, which is proportional to each drop's volume. The ratio of surface tension to the force of gravity on each drop is therefore proportional to the surface-area-to-volume ratio. For simplicity, I approximate the smaller drop as a sphere with diameter d. This means that I should compare the surface-area-to-volume ratio of this sphere with the surface-area-to-volume ratio of a sphere of diameter $10d$.

The surface area of a sphere of radius $R = \frac{1}{2}d$ is $4\pi R^2 = \pi d^2$, and its volume is $\frac{4}{3}\pi R^3 = \frac{4}{3}\pi(d/2)^3 = \frac{1}{6}\pi d^3$, and so the surface-area-to-volume ratio of a sphere is $(\pi d^2)/(\frac{1}{6}\pi d^3) = 6/d$. The surface-area-to-volume ratio thus decreases by a factor of 10 as the diameter of the drop increases by a factor of 10. ✔

Figure 18.27 A partially inflated balloon. The tension is greatest where the rubber is stretched the most.

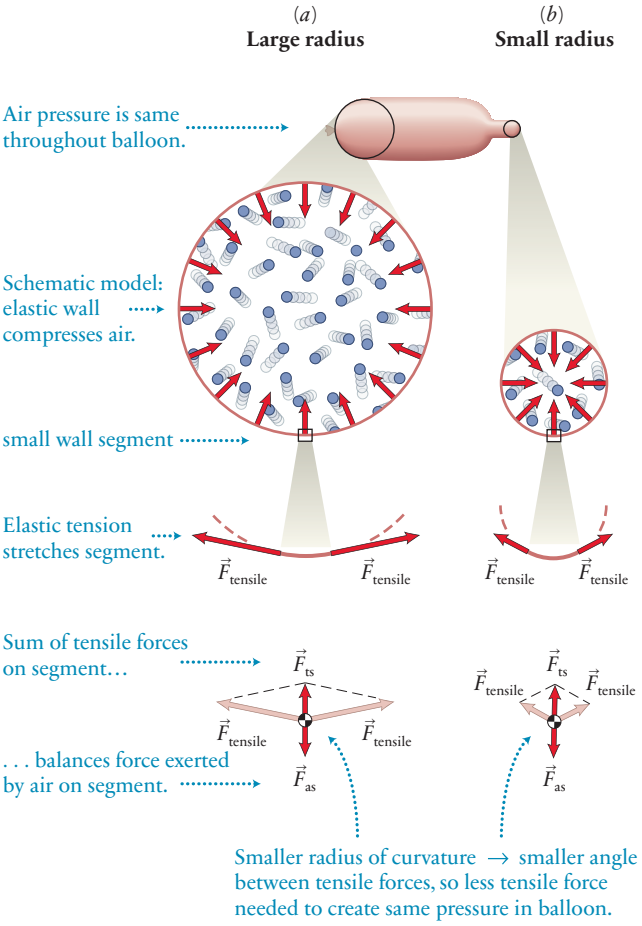


The analogy between surface tension and the tension in the rubber of a balloon allows us to draw some important conclusions. Consider, for example, the air-filled balloon of **Figure 18.27**. The rubber in the uninflated end of the balloon is less stretched than the rubber in the inflated part, and so there is less tension in the uninflated rubber than in the inflated rubber. (You may have noticed that balloons are softest and most easily squeezed where they are least inflated.) To understand how the tension in a balloon relates to its shape and to the pressure in the air inside the balloon, answer the following checkpoint first.

18.10 How does the air pressure in the uninflated part of the balloon in Figure 18.27 compare with the air pressure in the inflated part?

Did the answer to Checkpoint 18.10 surprise you? How can one part of the balloon be soft and another part hard if the air pressure is the same everywhere inside the balloon? The reason has to do with *radius of curvature*, which is the radius of the circle that approximates a curve at the point under consideration. As **Figure 18.28** shows, the radius of curvature is different in the two parts. The air pressure must be the same in the two parts, as you learned in Checkpoint 18.10, and the inward force exerted by the rubber surface of the balloon on the air inside is caused by the tension in the rubber. As you can see by comparing the free-body diagrams for a small surface segment in Figure 18.28*a* and *b*, the smaller the radius of curvature, the smaller the tensile force needed to provide a given inward force. In the part of the balloon where the radius of curvature is smaller, the angle of the tensile forces is large and so smaller tensile forces add up to the same vector sum. Indeed when the radius is smaller, the rubber is stretched less and so the tensile force is smaller.

Figure 18.28 The reason why the balloon in Figure 18.27 has an uninflated end. Because the air pressure must be the same everywhere inside a balloon, areas with a small radius of curvature must have smaller-magnitude tensile forces than areas with a larger radius of curvature.



CONCEPTS

In an elastic membrane enclosing a gas, tension in the membrane increases with increasing radius of curvature, but the pressure is the same everywhere inside the volume enclosed by the membrane, as required by Pascal's principle.

The situation is different with liquids because a liquid surface is not exactly like a stretched rubber skin. In particular, surface tension in a liquid is determined entirely by the cohesive forces between the liquid particles and is constant regardless of the shape of the liquid surface. So the tensile forces exerted on a given segment of the infinitesimally thin surface layer separating liquid and gas are the same no matter what the radius of curvature of the surface is. As shown in **Figure 18.29**, these constant tensile forces exerted on the surface segment of a liquid droplet cause a greater inward force when the radius of curvature is small.

Figure 18.29 Because surface tension in a liquid is independent of the shape of the liquid surface, a small liquid sphere has a greater internal pressure than a larger liquid sphere.

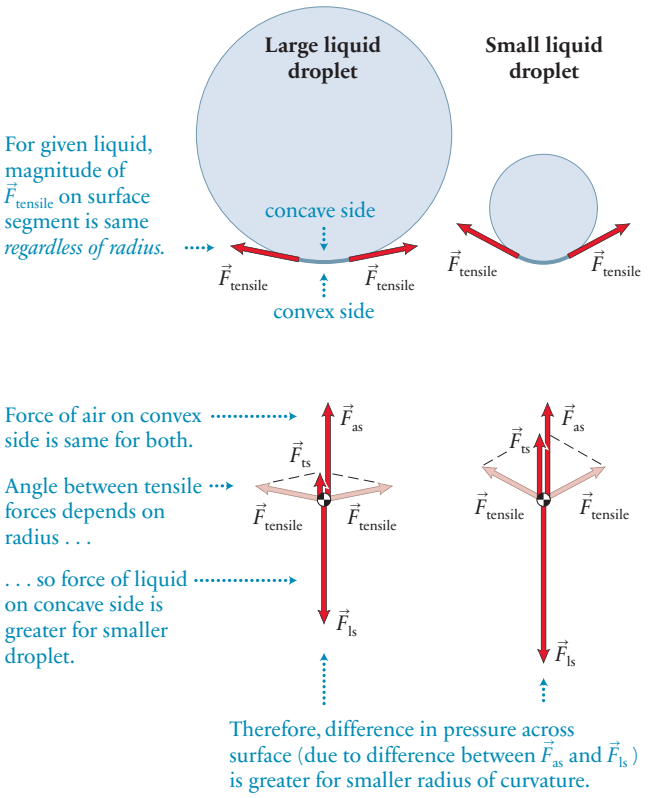
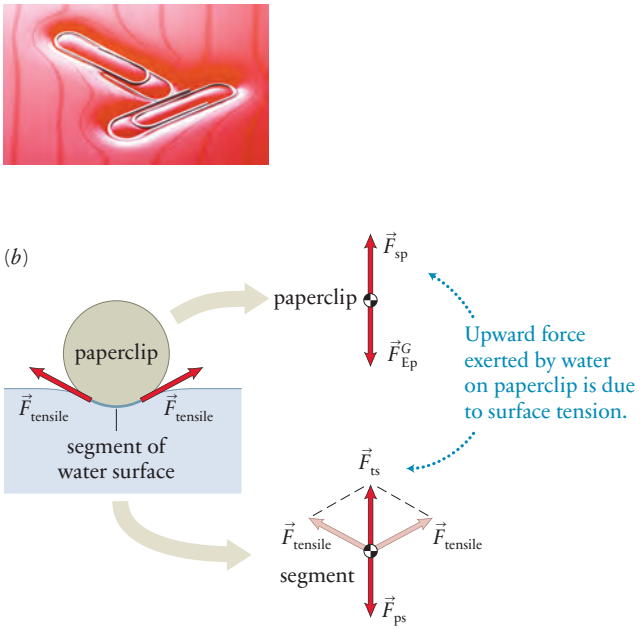


Figure 18.30 Surface tension enables paperclips to float on water.



Because the segment is at rest, this inward force on the segment must be balanced by the forces exerted by the liquid and air on either side of the surface layer. The outward force exerted by the liquid on the segment is therefore greater than the inward force exerted by the air on the segment. Because these forces result from the pressures in the fluids on either side of the surface, we can conclude that the pressure in the liquid on the concave side of the surface must be greater than the pressure in the air outside the droplet. We will refer to this difference as the *pressure difference across the surface:*

> **As the radius of curvature of a liquid surface decreases, the pressure difference across the surface increases, with the pressure being greatest on the concave side of the surface.**

This statement and the earlier boldfaced statement describing pressure in a membrane-enclosed gas are consequences of **Laplace's law,** which relates surface curvature to surface tension and to the pressure in a liquid.

Surface tension makes it possible for small insects to walk on a water surface and for small objects to float on a liquid surface. For example, if you carefully place a steel paperclip on a water surface, it stays there even though a thicker steel cylinder would sink (**Figure 18.30*a***). There are two ways to explain this. The first is to invoke Laplace's law: The water surface beneath the paperclip is concave up and so the pressure above the surface is greater than the pressure just below it. The pressure just below the surface where the paperclip rests is atmospheric pressure, while the pressure above the surface where the paperclip rests is atmospheric pressure plus the pressure caused by the force exerted by the paperclip on the surface.

The second way to explain why the paperclip floats is illustrated in Figure 18.30*b*, which shows free-body diagrams for the segment of the surface just below the paperclip and for the paperclip. The tensile forces due to surface tension balance the downward force $\vec{F}_{\text{ns}}$ exerted by the paperclip on the surface segment. This downward force $\vec{F}_{\text{ns}}$ forms an interaction pair with the upward force $\vec{F}_{\text{sn}}$ exerted by the surface segment on the paperclip, which balances the downward force of gravity on the paperclip and makes the paperclip float.

18.11 (*a*) What is the difference between the pressure just below the water surface in a pool and the pressure in the air just above it? (*b*) Is the pressure inside a raindrop greater than, equal to, or smaller than the pressure in the surrounding air? (*c*) How does the pressure inside a small raindrop compare with the pressure inside a larger raindrop?

CONCEPTS

Figure 18.31 Small drops are nearly spherical; larger ones flatten out. For a given combination of liquid and solid surface, the contact angle is fixed and does not depend on the drop size or shape.

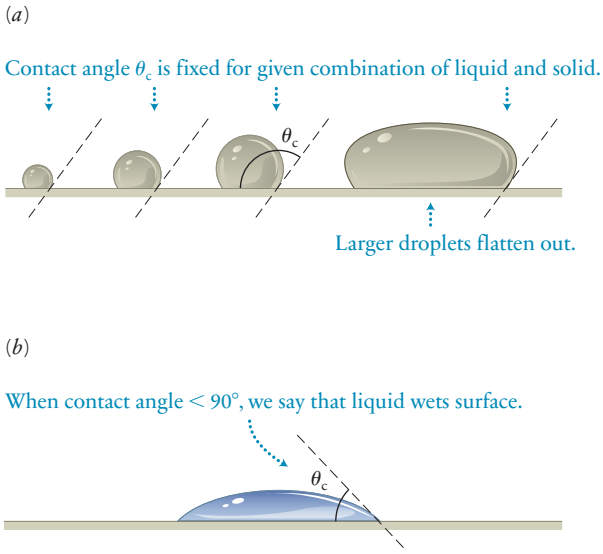
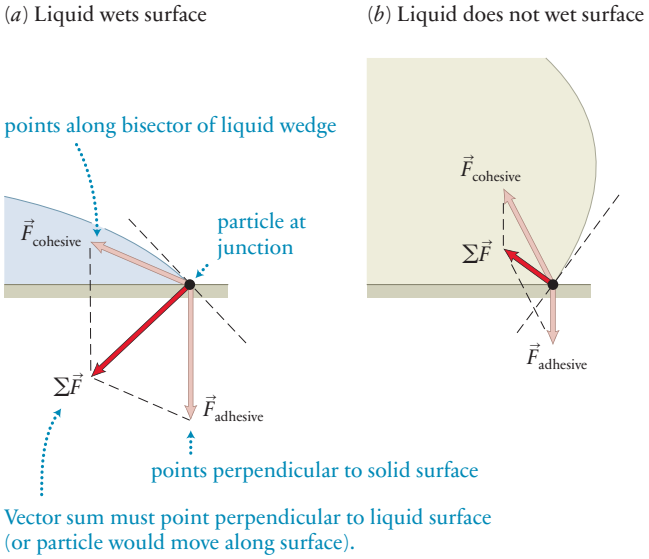


Figure 18.32 Free-body diagrams for a particle at the gas-liquid-solid junction for a liquid that (*a*) wets and (*b*) does not wet a horizontal surface.



A liquid surface that comes in contact with a solid surface forms an angle with the solid surface (**Figure 18.31**). This angle, called the *contact angle*, is the same for a given liquid-solid combination (Figure 18.31*a*).* The contact angle is defined as the angle between the solid surface and the tangent to the liquid surface at the point where liquid and solid meet; the contact angle is always measured in the liquid. If the contact angle is smaller than 90°, the liquid is said to *wet* the solid surface (Figure 18.31*b*). If the contact angle is greater than 90°, as in Figure 18.31*a*, the liquid does not wet the surface; it forms drops instead of spreading out and wetting the surface.

The contact angle is strongly influenced by how smooth and clean the solid surface is and by the presence of impurities in the liquid. For example, water beads up into drops on a paraffin wax surface, but adding a bit of soap to the water makes the drop spread out and wet the paraffin surface.

The contact angle for a liquid on a solid surface is due to the interplay between the cohesive forces between the particles in the liquid and the *adhesive forces*† between the liquid particles and the atoms in the solid. Consider, for example, a particle at the point where the surface meets the solid (**Figure 18.32**). This particle is subject to a cohesive force exerted by a wedge of liquid and an adhesive force exerted by the solid surface. Symmetry requires the cohesive force to be directed along the bisector of the wedge. Likewise, the adhesive force must be perpendicular to the solid surface. (If the adhesive force were not perpendicular to the solid surface, there would be a flow tangential to the surface in response to any tangential component of that force.) Ignoring the effects of gravity (which are very small on a single particle), we can say that the vector sum of the forces exerted on the particle is the sum of the cohesive force and the adhesive force. Because the liquid cannot sustain any shear force, the vector sum of the forces on any particle at its surface must always be perpendicular to the surface (see Section 18.1). Given the directions of these two forces and their vector sum, their relative magnitudes are fixed. For example, for a liquid that wets a solid with a contact angle of 60° or smaller, the adhesive force is always greater than the cohesive force (Figure 18.32*a*), while for a nonwetting combination, the cohesive force is always greater than the adhesive force (Figure 18.32*b*).

The contact angle also determines the shape of the edges of a liquid surface at any container walls (**Figure 18.33**). In a very narrow tube, called a *capillary* tube (after the Latin word for *hair*), the surface of a liquid is curved. This curved surface, called a **meniscus,** is concave when the liquid wets the walls of the capillary tube (Figure 18.33*a*) and convex when the liquid is nonwetting (Figure 18.33*b*). By looking closely at the edge of the surface of a liquid in a glass, you can determine whether or not the liquid wets the glass.

CONCEPTS

*The gas surrounding the liquid and solid also affects the contact angle. Because this effect is generally very small, we shall ignore it.

†Cohesive forces are exerted between like particles. Adhesive forces are exerted between unlike particles.

Figure 18.33 The concave or convex shape of the surface of a liquid in a capillary tube is determined by the contact angle.

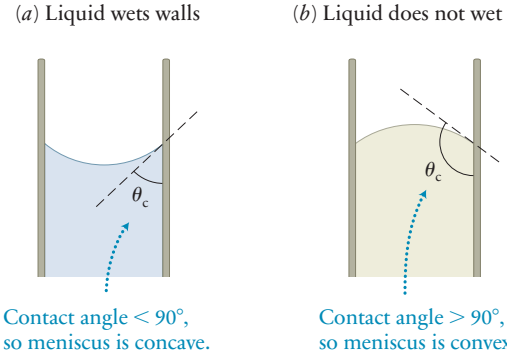
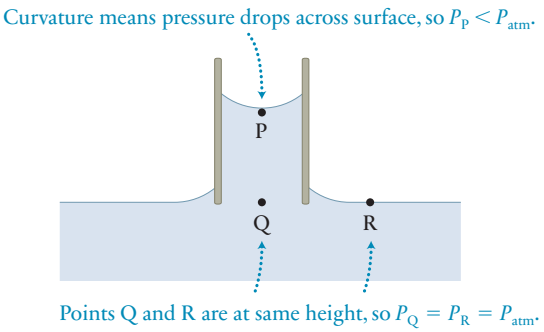


Figure 18.35 The pressure difference across the concave surface of a liquid that wets the walls of a capillary tube causes the liquid to rise in the tube, a phenomenon called *capillary rise*.



The free-body diagrams for a particle at the point where the liquid surface meets the glass are shown in **Figure 18.34** for a concave meniscus and a convex meniscus. Notice the similarity with the free-body diagrams in Figure 18.32.

The formation of a meniscus is responsible for the rise of liquid in a capillary tube, as **Figure 18.35** shows. Because the meniscus is concave, the pressure at point P in the liquid, just below the meniscus, must be smaller than the atmospheric pressure just above the meniscus. The pressure at Q, a point directly below P and at the same height as the surface of the liquid surrounding the capillary tube, must be the same as the pressure at R (otherwise the liquid would flow so as to eliminate any horizontal pressure difference). Because the surface at R is flat, there is no pressure difference across the surface, and so we know that the pressure at R (and therefore the pressure at Q) is equal to atmospheric pressure. The pressure difference between points P and Q causes an upward buoyant force to be exerted by the liquid in the tube between P and Q, causing a *capillary rise* until the downward force of gravity exerted on the liquid that has been raised into the tube balances the upward force due to the pressure difference. As illustrated in **Figure 18.36**, capillary rise is more pronounced in a narrow tube. The narrower the tube, the smaller the radius of curvature of the meniscus. As we know from Laplace's law, a smaller radius of curvature means a greater pressure difference across the surface and therefore a greater upward force on the liquid.

18.12 Explain why surface effects cannot support a large solid cylinder made of the same material as a needle.

Figure 18.34 Free-body diagrams for a particle at the gas-liquid-solid junction for a liquid that (*a*) wets and (*b*) does not wet a vertical surface.

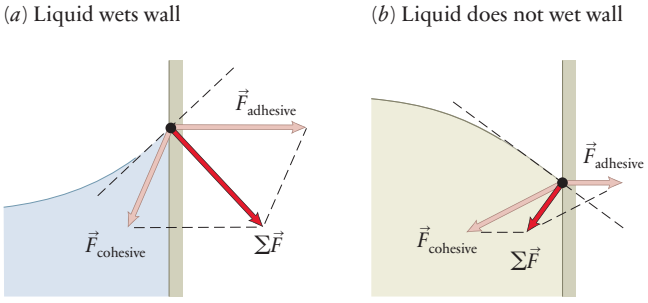
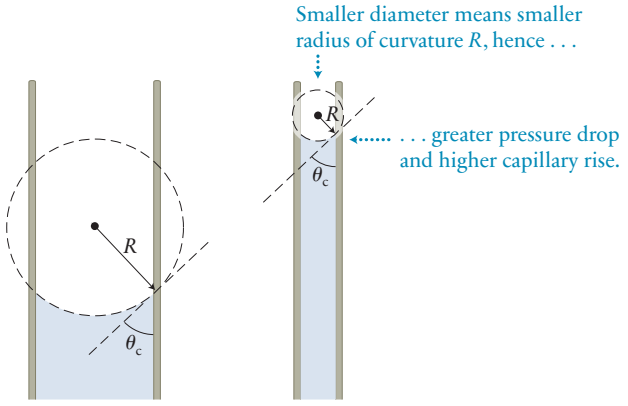


Figure 18.36 The narrower the tube, the greater the capillary rise.

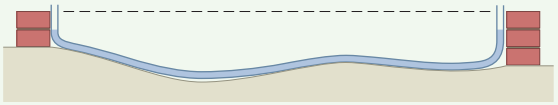


CONCEPTS

Self-quiz

1. Masons use a hose open at both ends and filled with water to ensure that the blocks they lay remain level (**Figure 18.37**). How does such a hose help them do this?

Figure 18.37

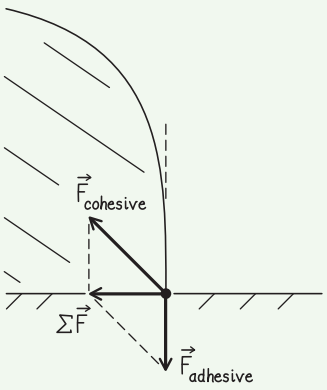


2. (*a*) Imagine a rock and a balloon of the same volume completely submerged in water. How do the buoyant forces exerted on them compare? (*b*) How do the buoyant forces compare when the balloon volume is half the rock volume? (*c*) Which object floats and which sinks in part *a* and in part *b*?
3. For a thrown ball to rise, what type of spin should the thrower place on it?
4. (*a*) Draw a free-body diagram analogous to Figure 18.32 for a gas-liquid-solid boundary at which the contact angle is 90°. Make your solid surface horizontal. (*b*) How do the magnitudes and directions of the adhesive and cohesive forces exerted on your particle compare? (*c*) How does the magnitude of the cohesive force between two liquid particles compare with the magnitude of the adhesive force between a liquid particle and an atom in the solid?
5. (*a*) Imagine a capillary tube placed in a nonwetting liquid (as opposed to the wetting-liquid situation of Figure 18.35). How does the pressure at a point P just below the now convex surface compare with atmospheric pressure? (*b*) How does the level in the tube compare with the level of the liquid outside the tube?

Answers

1. The air pressure at both water surfaces is atmospheric pressure. Therefore the height of the water surface at the left end of the hose must be the same as the height of the water surface at the right end, in order to balance the pressure throughout the water. If the masons align the two ends of the hose with the two ends of a row of blocks and no water bubbles out of either end of the hose, the two ends of the wall are at the same height, meaning the row is level.
2. (*a*) Because the rock and balloon occupy the same volume, they displace the same volume of water, and so the surrounding water exerts equal buoyant forces on them. (*b*) The smaller volume of the balloon displaces a smaller volume of water, and so the buoyant force exerted on the balloon is half the buoyant force exerted on the rock. (*c*) In either case, the balloon rises and the rock sinks. The buoyant force on a rock is always smaller than the force of gravity exerted on it. Likewise the buoyant force on a submerged balloon is always greater than the force of gravity exerted on it, regardless of the volume of the balloon.
3. The thrower needs to place a "backspin" on the ball, which is a spin in which the top of the ball moves toward the thrower. This direction of spin causes the speed of the air relative to the surface of the ball to be greater at the top of the ball than at the bottom. Consequently the air pressure is smaller at the top of the ball than at the bottom. As a result, the ball is subject to an upward force due to this pressure difference.
4. (*a*) See **Figure 18.38**. A contact angle of 90° means that the tangent line to the drop at the contact point is vertical. (*b*) The cohesive force makes an angle of 45° with the horizontal, and the adhesive force is directed vertically downward. The vector sum of the forces must be horizontal, perpendicular to the tangent to the drop at the contact point. The cohesive force is greater than the adhesive force because the vertical component of the cohesive force must cancel the vertically downward adhesive force. (*c*) They are equal. For a particle on the liquid-solid boundary, there are solid atoms below the boundary, liquid particles above the boundary, and liquid particles to the left and the right. The vector sum of the forces exerted on the liquid particle at the boundary is zero and so the magnitude of the cohesive forces between two liquid particles must be the same as the magnitude of the adhesive force between a liquid particle and an atom in the solid.

Figure 18.38



5. (*a*) Because the surface is convex, the pressure at P is greater than atmospheric pressure. (*b*) Because the pressure at R is still atmospheric pressure, the difference in pressure causes a downward force on the liquid in the tube and a resulting *capillary fall*.

18.5 Pressure and gravity

In the first part of this chapter we defined the ratio of the magnitude of the force $\vec{F}^{\perp}_{\text{fs}}$ exerted by a fluid on a surface and the area A of that surface as the **pressure** in a fluid:

$$P \equiv \frac{F^{\perp}_{\text{fs}}}{A}. \tag{18.1}$$

Pressure is a scalar, and the SI unit of pressure is newtons per meter-squared, N/m^2. This derived unit is called the **pascal** (Pa), in honor of the French scientist Blaise Pascal:

$$1\ \text{Pa} \equiv 1\ \text{N/m}^2. \tag{18.2}$$

This definition means that a fluid in which the pressure is 1 Pa exerts a force of magnitude 1 N on a 1-m^2 area of the container that holds the fluid. Because a pressure of 1 Pa is very small, pressures are often given in kilopascals: 1 kPa = 1000 Pa. For example, the average pressure in the atmosphere at sea level is

$$P_{\text{atm}} = 101.3\ \text{kPa}. \tag{18.3}$$

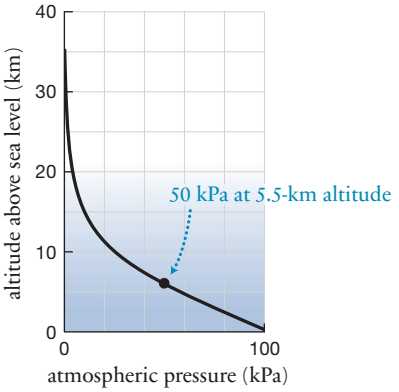
This pressure is sometimes referred to as 1 atmosphere (1 atm). *Atmosphere* is not an SI unit, however, and so any pressure given in atmospheres must be converted to an SI unit (pascals or kilopascals) before you can do most calculations.

Inside a box that contains absolutely no atoms or molecules, the pressure is zero because nothing exerts a force on the walls of the box. The box is said to contain a *perfect vacuum*. In practice, a perfect vacuum does not exist because it is impossible to remove every particle from a volume of space. Even in deep outer space, the pressure, though very small, is nonzero. In practice, a *vacuum* is any space in which the pressure is less than about 1 Pa. In such a vacuum, $P = 0$ is a very good approximation.

There are several contributions to pressure in a fluid. One contribution is from the gravitational force exerted on the fluid. The pressure in the atmosphere at sea level, for example, is due to the gravitational force exerted by Earth on all the air surrounding the planet. Other contributions are from the motion of the fluid and, when the fluid is a liquid, from the curvature of the liquid surface. A fourth contribution is from the collisions that take place between the fluid particles, and between the fluid particles and the container walls—the pressure in an inflated balloon is an example of this fourth contribution. We'll discuss this contribution to the pressure in a gas in more detail in Chapter 19. In the remainder of this chapter, we concentrate on the gravitational, motion, and surface contributions to pressure.

Let us begin by looking at a familiar type of pressure, the *atmospheric pressure* that results from the air surrounding Earth. The atmosphere can be modeled as an "ocean" of air extending from the ground up. Because gases are compressible, the distribution of gas molecules in the air is not uniform. All the molecules are attracted to the ground by Earth's gravitational pull, which means that higher-up molecules press down on lower ones, compressing the air and making the mass density high near sea level and lower at altitudes above sea level. Because gases are compressible, the relationship between atmospheric pressure and altitude is complicated by the fact that the mass density of air varies with altitude. **Figure 18.39** shows how the average atmospheric pressure varies with altitude. If the mass density of air were constant, the curve would be a straight line with a negative slope. If the extremes were as in Figure 18.39—100 kPa at sea level and essentially zero at 30 km—the atmospheric pressure at the midpoint

Figure 18.39 Variation in atmospheric pressure with altitude above sea level.



QUANTITATIVE TOOLS

between these two altitudes, 15 km, would be half of its sea-level value. Instead we see a nonlinear relationship in Figure 18.39, with the pressure at 50 kPa, half the sea-level value, not at an altitude of 15 km but at a mere 5.5 km. This means that half of the atmosphere is below 5.5 km. At the cruising altitude of a commercial airplane (10 km), the atmospheric pressure is about one-quarter of what it is at sea level, meaning that 75% of the atmosphere is below the airplane. At an altitude of 100 km, the pressure in the atmosphere drops to about 1 Pa, close to vacuum.

Even though the changes in atmospheric pressure are considerable over large distances, they are usually negligible over small vertical distances because, even at its densest, at sea level, the mass density of air is so small. Near sea level, for example, the atmospheric pressure drops 12 Pa for every meter of altitude. The pressure difference between the air near your feet and the air near your head is therefore about 20 Pa, which is negligibly small relative to atmospheric pressure, 1×10^5 Pa.

EXAMPLE 18.5 Altimeter

Airplane cabins are pressurized to prevent passengers from developing altitude sickness during flight. Your altimeter watch, which uses pressure to determine altitude, indicates a cabin altitude of 2500 m even though the captain has just announced that the airplane is flying at an altitude of 10,000 m. If you ignore the motion of the airplane, what do you calculate to be the integral of the magnitude of the force exerted by the air on the fuselage of the airplane, which you estimate to be a cylinder of length 60 m and diameter 6.0 m?

❶ **GETTING STARTED** Pressure decreases with altitude as shown in Figure 18.39. Because the cabin is pressurized, the pressure inside the cabin is greater than it would be if it were allowed to equalize with the exterior atmospheric pressure at 10,000 m. That the altimeter indicates an altitude of 2500 m tells me that the cabin pressure is equal not to atmospheric pressure at sea level (about 100 kPa) but to a value somewhat lower than 100 kPa. (Because I can feel my ears "pop" as the airplane takes off and lands, I know that the cabin pressure must indeed change somewhat.)

❷ **DEVISE PLAN** The force exerted by the air on the fuselage is the vector sum of two forces: an outward force due to the pressure P_{in} inside the cabin and an inward force due to the pressure P_{out} outside the cabin. According to Eq. 18.1, the magnitude of each of these forces is PdA, where dA is the surface area of a segment of the fuselage. Because $P_{in} > P_{out}$, the air exerts on each fuselage segment an outward force of magnitude $(P_{in} - P_{out})dA$. Figure 18.39 gives me the values for P_{out} at 10,000 m and for what the cabin pressure P_{in} must be given that it corresponds to atmospheric pressure at 2500 m. Ignoring the tiny variation in air pressure over the height of the airplane, the values of P_{in} and P_{out} are identical over the entire fuselage, and so I just need to multiply the pressure difference by the surface area $A = \int dA$ of the entire fuselage to compute the integral of the magnitude of the forces exerted on the fuselage.

❸ **EXECUTE PLAN** From Figure 18.39, I see that $P_{out} = P_{10\,km} = 24$ kPa and $P_{in} = P_{2.5\,km} = 80$ kPa, so $P_{in} - P_{out} = 80$ kPa $-$ 24 kPa $= 56$ kPa. The surface area of the (assumed) cylindrical fuselage is equal to circumference times length plus the surface area of the two circular ends: $A = 2\pi R\ell + 2\pi R^2 = 2\pi(3.0\text{ m})(60\text{ m}) + 2\pi(3.0\text{ m})^2 = 1.2 \times 10^3\text{ m}^2$. The sum of the magnitudes of the forces exerted by the air on the fuselage is thus $(56\text{ kPa})(1000\text{ Pa}/1\text{ kPa})(1.2 \times 10^3\text{ m}^2) = 6.7 \times 10^7$ N. ✔

❹ **EVALUATE RESULT** This force magnitude is very large, corresponding to the force of gravity exerted on 6.8×10^6 kg, or 6800 metric tons! The fuselage does not rupture because the force is distributed over a very large area. For comparison, if the pressure inside the cabin were maintained at a pressure equal to the atmospheric pressure at sea level, the pressure difference across the fuselage would be $P_{in} - P_{out} = 101$ kPa $-$ 24 kPa $= 77$ kPa and the force would be some 40% greater.



18.13 Consider a book lying on a table at sea level with the front cover facing up. The book is 0.28 m tall, 0.22 m wide, and 50 mm thick; its mass is 3.0 kg. How does the force exerted by the atmosphere on the front cover compare with the force of gravity exerted on the book?

The answer to Checkpoint 18.13 is astonishing. With the book under so much pressure, how can anyone lift it? The reason you can is that the air is pushing on the book from all directions. Not only does it push down on the front cover, it also pushes up on the bottom cover. Because the surface areas of the top and bottom covers are the same and because the air

pressure below the book is virtually the same as that above the book, the magnitudes of the upward and downward forces exerted by the atmosphere on the book are the same, and so their vector sum is zero.

You may wonder how the atmosphere can exert an upward force on the bottom cover if the book is lying flat on a table. The reason is that there is always a tiny bit of air between the cover and the table surface, and the air pressure in this layer has to be the same as the air pressure everywhere else around the book (otherwise air would rush into or out of the space between table and cover to eliminate the pressure difference). If you were to remove the air from underneath the book—say, by putting the book on top of the opening of a chamber from which the air has been evacuated—it would be nearly impossible to lift the book off the table.



18.14 A dart that has a suction cup at one end sticks to a ceiling. Describe what force holds the dart against the ceiling.

In contrast to gases, liquids are nearly incompressible, which means that we can take the mass density in a liquid to be uniform. Therefore we can derive a simple expression to describe how the pressure in a liquid varies with height. Consider a cylindrical volume of liquid inside a large stationary body of identical liquid (**Figure 18.40***a*). The cylindrical volume has volume V and horizontal cross-sectional area A, and Figure 18.40*b* shows that three forces are exerted on it in the vertical direction: an upward force exerted by the column of liquid below the bottom of the cylindrical volume (point 1), a downward force exerted by the column of liquid above the top of the volume (point 2), and a downward force of gravity. Because our volume is at rest, the vector sum of the three forces must be zero, and so, with the positive y axis pointing upward, we can say for all the forces exerted on the cylindrical volume in the vertical direction,

$$\Sigma F_{cy} = F^{c}_{1cy} + F^{c}_{2cy} + F^{G}_{Ecy} = P_1A - P_2A - mg = 0, \qquad (18.4)$$

where m is the mass of the liquid in the cylindrical volume. Because the mass density ρ of the liquid is the same everywhere, we can write for the mass of the cylindrical volume $m = \rho V = \rho A(y_2 - y_1)$, and so

$$P_1A - P_2A - \rho A(y_2 - y_1)g = 0 \qquad (18.5)$$

or, dividing all terms by A and bringing terms with the same subscripts to the same side,

$$P_1 + \rho g y_1 = P_2 + \rho g y_2 \quad \text{(stationary liquid)}. \qquad (18.6)$$

Equation 18.6 tells us that the quantity $P + \rho gy$ is the same at the top and bottom of the cylindrical volume. Because the top and bottom are at arbitrary positions, the fact that $(P + \rho gy)_{top} = (P + \rho gy)_{bottom}$ means that the quantity $P + \rho gy$ is the same *everywhere* in the *entire* liquid. In other words, the pressure in a stationary liquid is the same at all points in any given horizontal plane through the liquid (that is, at some fixed value of y).

Rearranging Eq. 18.6 to bring all the terms containing y to one side ($P_1 = P_2 + \rho gy_2 - \rho gy_1$), we see that the pressure at all points in a horizontal plane 1 located a vertical distance $d = y_2 - y_1$ below another horizontal plane 2 is given by

$$P_1 = P_2 + \rho g d \quad \text{(stationary liquid 1 below 2)}. \qquad (18.7)$$

Figure 18.40 Model for describing how the pressure in a liquid varies with height.

(*a*) Imaginary cylindrical volume of liquid at rest inside larger volume of same liquid (*b*) Free-body diagram for cylinder

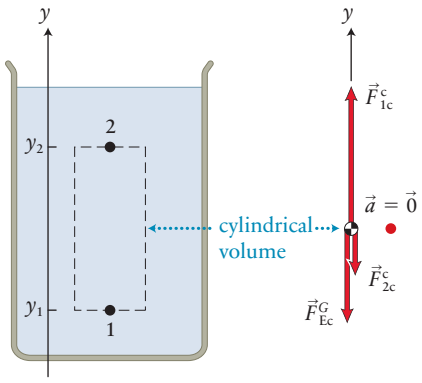


Figure 18.41 If the cylindrical volume is defined such that its top is at the liquid surface, the pressure at the top is equal to atmospheric pressure.

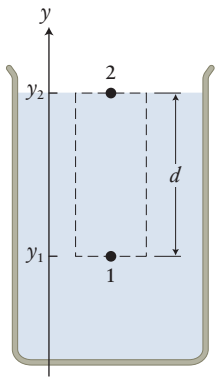
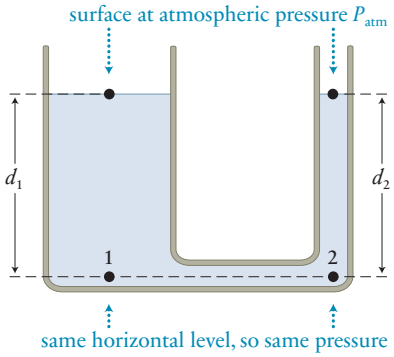


Figure 18.42 Liquid rises to the same level in connected tubes, regardless of their width or shape.



If we let point 2 be at the surface of the liquid (**Figure 18.41**), P_2 is the pressure at the surface, and Eq. 18.7 tells us that the pressure at a depth $d = y_2 - y_1$ below the surface is given by

$$P = P_{surface} + \rho g d \quad \text{(stationary liquid)}. \tag{18.8}$$

This expression shows that the pressure P in a stationary liquid increases with depth or, put differently, P decreases linearly with the height in the liquid, just as we concluded in Section 18.2.

The term that contains g is the gravitational contribution to the pressure and is sometimes referred to as *hydrostatic pressure*.

Now consider two liquid-filled tubes, one wide and one narrow and both open at one end, connected as shown in **Figure 18.42**. In each tube, the pressure at the surface of the liquid is equal to atmospheric pressure. This means that the downward force $P_{atm}A_{wide}$ exerted by the atmosphere on the surface of the liquid in the wide tube is significantly greater than the downward force $P_{atm}A_{narrow}$ exerted by the atmosphere on the surface of the liquid in the narrow tube. How does this difference in force affect the levels of the water in each tube?

To answer this question, let's consider the pressures at points 1 and 2 in Figure 18.42. Because these points are in the same horizontal plane, the pressures must be equal, $P_1 = P_2$. Equation 18.8 tells us that these pressures are related to the (atmospheric) pressure at the surface by

$$P_1 = P_{atm} + \rho g d_1 \quad \text{and} \quad P_2 = P_{atm} + \rho g d_2. \tag{18.9}$$

Because $P_1 = P_2$, we must have $d_1 = d_2$: The liquid rises to the same height in both tubes, even though the forces exerted by the atmosphere on the surfaces of the liquid are different. The reason is that, despite the difference in these downward forces, the *pressure* at the bottom of each tube must be the same ($P_1 = P_2$). If the pressures were not the same, the pressure difference would cause the liquid to flow until the pressure difference is zero, which occurs only when $d_1 = d_2$. The equal height of the liquid surfaces in the two tubes means that these surfaces are in the same horizontal plane. So the pressures at two points in a horizontal plane are the same as long as the two points are connected by an uninterrupted path through the liquid.

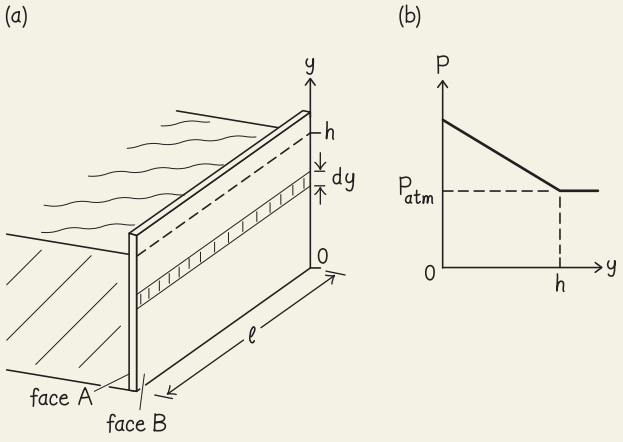
EXAMPLE 18.6 Dam

A dam of horizontal length ℓ holds water of mass density ρ to a height h. What is the magnitude of the force exerted by the water on the dam?

❶ **GETTING STARTED** I begin with a sketch (**Figure 18.43*a***). The pressure in the water behind the dam is equal to atmospheric pressure at the water surface and increases linearly with depth as given by Eq. 18.8. If I choose my y axis pointing upward, with the origin at the bottom of the dam, the pressure varies with y as shown in Figure 18.43*b*. For convenience, I call the water side of the dam face A and the opposite side face B.

❷ **DEVISE PLAN** Because the pressure varies with y, the magnitude of the force $\vec{F}_{wd}$ also varies with y. I therefore divide the area of face A into thin horizontal strips each of width dy (Figure 18.43*a*). Because I know the pressure at a given value of y, I can calculate the force dF_{wd} exerted on the strip. To calculate the force exerted by the water on the whole dam, I integrate my expression over y from $y = 0$ at the bottom of the dam to $y = h$ at the water surface.

Figure 18.43



QUANTITATIVE TOOLS

3 EXECUTE PLAN Because a point in the water located at some arbitrary value of y is at a depth $d = h - y$ below the water surface and because $P_{\text{surface}} = P_{\text{atm}}$, I write Eq. 18.8 in the form

$$P(y) = P_{\text{atm}} + \rho g(h - y).$$

The area of each strip is $dA = \ell\, dy$, and so the magnitude of the force exerted by the water on each strip is $dF = P(y)dA = P(y)\ell\, dy$. The magnitude of the force exerted by the water on the entire dam is then

$$F_{\text{wd}} = \int dF_{\text{ws}} = \ell \int_0^h P(y)dy = \ell\,[P_{\text{atm}}y + \rho g h y - \tfrac{1}{2}\rho g y^2]_0^h$$
$$= \ell h\,(P_{\text{atm}} + \tfrac{1}{2}\rho g h). \checkmark$$

4 EVALUATE RESULT The product ℓh is the surface area of the dam, and so the term $\ell h P_{\text{atm}}$ is the magnitude of the force exerted by the atmosphere on the dam. Because the atmosphere exerts a force of the same magnitude on face B in the opposite direction, the first term in my answer drops out of the vector sum of the forces exerted by the water on face A of the dam and by the atmosphere on face B of the dam. The remaining term, $\ell h(\frac{1}{2}\rho g h)$, is the product of the area of the face of the dam and $\frac{1}{2}\rho g h$, which is the pressure increase halfway down in the water. Because the pressure increases linearly with depth, this term represents the average value of the pressure increase between surface and bottom, and so my answer makes sense.

Let us now turn to buoyancy. The volume of liquid in Figure 18.40 is held in place by the buoyant force $\vec{F}^{\text{b}}_{\text{lv}}$ exerted by the surrounding liquid. The vector sum of the forces exerted on the volume of liquid is thus

$$\Sigma\vec{F}_{\text{v}} = \vec{F}^{\text{b}}_{\text{lv}} + \vec{F}^{G}_{\text{Ev}} = \vec{0}. \tag{18.10}$$

Because the magnitude of the force of gravity on the volume of liquid is $\rho V g$, we obtain for the magnitude of the **buoyant force** exerted by the surrounding liquid on the volume of liquid

$$F^{\text{b}}_{\text{lv}} = \rho V g. \tag{18.11}$$

The direction of the buoyant force is always opposite the direction of the gravitational acceleration. Because this expression holds for any object in any fluid, we can write for the buoyant force on any object o in a fluid f:

$$F^{\text{b}}_{\text{fo}} = \rho_{\text{f}} V_{\text{disp}} g, \tag{18.12}$$

where we have added a subscript to V to remind us that this volume represents the *displaced volume* of fluid and a subscript on ρ to remind us that it is the mass density of the fluid.

The vector sum of the upward buoyant force and the downward gravitational force exerted on an object in a given fluid determines whether the object sinks or floats in that fluid. When they are equal in magnitude, $F^{\text{b}}_{\text{fo}} = F^{G}_{\text{Eo}}$ or $\rho_{\text{f}} V_{\text{disp}} g = mg = \rho_{\text{o,av}} V_{\text{o}} g$, where we have written the mass of the object as the product of its average mass density $\rho_{\text{o,av}}$ and volume V_{o}. For an object that is floating, the displaced volume is less than the volume of that object:

$$V_{\text{disp}} = \frac{\rho_{\text{o,av}}}{\rho_{\text{f}}} V_{\text{o}} \quad \text{or} \quad \frac{\rho_{\text{o,av}}}{\rho_{\text{f}}} < 1. \tag{18.13}$$

Therefore we have

$$\rho_{o,\text{av}} < \rho_{\text{f}} \quad \text{(object floats)}. \tag{18.14}$$

An object placed in a fluid floats whenever the average mass density of the object is smaller than the mass density of the fluid.

Conversely, an object sinks when the magnitude of the buoyant force is smaller than that of the gravitational force. Working through the same thought sequence that got us from forces to Eq. 18.14, we see that

$$\rho_{o,av} > \rho_f \quad \text{(object sinks)}. \tag{18.15}$$

An object placed in a fluid sinks whenever the average mass density of the object exceeds the mass density of the fluid.

When the object's average mass density is equal to the mass density of the fluid, $\rho_{o,av} = \rho_f$, the object is said to have *neutral buoyancy*—it can "hang" motionless anywhere in the fluid.

Figure 18.44 Simple mechanical pressure gauge.



18.15 An object floats with 80% of its volume submerged in water. How does the average mass density of the object compare with that of water?

18.6 Working with pressure

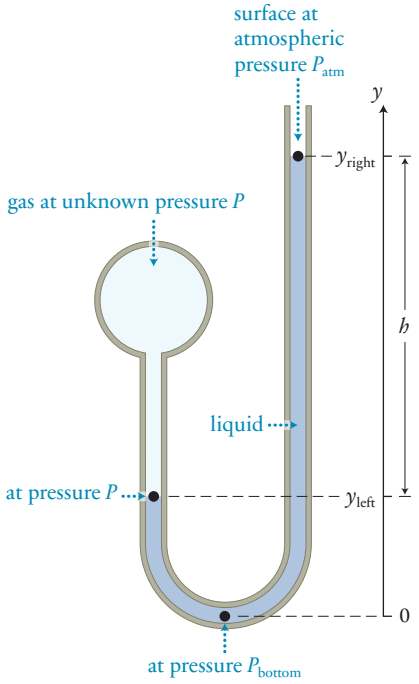
The results in Section 18.5 allow us to understand a number of applications of pressure. Let's first consider measurements of pressure. A *pressure gauge* is used to measure pressures that are near or greater than atmospheric pressure; a *vacuum gauge* is used to measure pressures that are smaller than atmospheric pressure.

The most common pressure gauges have a flexible, sealed chamber that changes shape when the pressure inside or outside the chamber changes. The deformation of the chamber is measured electrically, optically, or mechanically. For example, in a simple mechanical pressure gauge (**Figure 18.44**), the chamber deformation is coupled via a gear to an indicating needle on a calibrated dial. The reading on a pressure gauge is the pressure in excess of atmospheric pressure. Thus when you use a pressure gauge to check the pressure in, say, a bicycle tire, the reading on the gauge, called *gauge pressure* P_{gauge}, is related to the actual pressure P in the tire by

$$P = P_{gauge} + P_{atm}. \tag{18.16}$$

For example, when a tire pressure gauge reads 2.4 atm, the pressure inside the tire is 3.4 atm. If the pressure inside a container is smaller than atmospheric pressure, the pressure gauge reading is negative.

Figure 18.45 An open-tube manometer. The difference in the levels in the two legs of the manometer gives the pressure P inside the bulb.



The simplest pressure gauge is the open-tube *manometer* (**Figure 18.45**). A manometer is essentially a liquid-filled, U-shaped tube, one leg of which is connected to the container where the pressure P is to be measured and the other leg of which is open to the atmosphere. The pressure in the liquid at the bottom of the U is P_{bottom}, while that at the surface of the liquid in the right leg is atmospheric pressure. Applying Eq. 18.7 to these two points, we get $P_{bottom} = P_{atm} + \rho g y_{right}$. Applying the same procedure to the left leg, we have $P_{bottom} = P + \rho g y_{left}$, and so $P = P_{atm} + \rho g(y_{right} - y_{left})$ or, if we denote the difference in the heights of the liquid in the two legs by h, the pressure measured by the manometer is

$$P = P_{atm} + \rho g h. \tag{18.17}$$

Equation 18.17 shows that using a manometer to measure the pressure in some fluid-filled container requires knowing atmospheric pressure, which means we cannot use a manometer to measure atmospheric pressure. We can use another device, though, a *barometer*, illustrated in

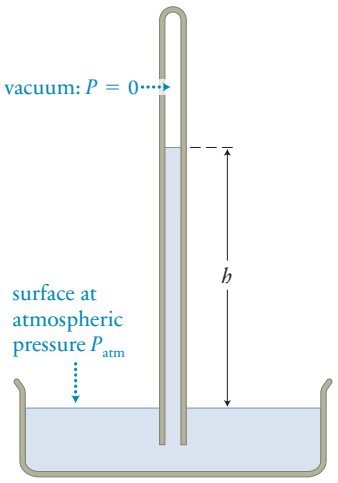
QUANTITATIVE TOOLS

Figure 18.46. This device consists of a shallow reservoir; a long, thin tube closed at one end and open at the other; and a dense liquid, usually mercury. Some of the liquid sits in the reservoir. The tube is filled completely with the liquid, and its open end is covered temporarily. The tube is then inverted and its open end plunged into the liquid in the reservoir. As the temporary cover is removed, some of the liquid drains from the tube into the reservoir, leaving a vacuum above the liquid in the closed end of the tube. The level of the liquid in the tube remains above the level of the liquid in the reservoir so as to satisfy Eq. 18.7 when we define the liquid surface in the reservoir as point 1 and the liquid surface in the tube as point 2. When we let the reservoir liquid surface be at $y_1 = 0$, the left side of Eq. 18.6 becomes $P_{atm} + 0$. Because there is a vacuum above the tube liquid surface, the right side of Eq. 18.6 is $0 + \rho gh$, and so we have

$$P_{atm} = \rho gh. \tag{18.18}$$

If we know the mass density ρ of the liquid, we can determine atmospheric pressure by measuring the height h of the liquid in the tube.

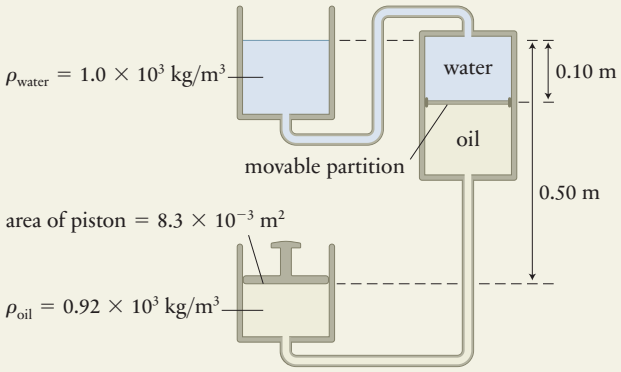
Figure 18.46 A barometer. The height of the column of liquid in the tube gives the atmospheric pressure.



EXAMPLE 18.7 Liquid support

In **Figure 18.47**, the top tank, which is open to the atmosphere, contains water ($\rho_{water} = 1.0 \times 10^3$ kg/m^3) and the bottom tank contains oil ($\rho_{oil} = 0.92 \times 10^3$ kg/m^3) covered by a piston. The tank on the right has a freely movable partition that keeps the oil and water separate. The partition is a vertical distance 0.10 m below the open surface of the water. If the piston in the bottom tank is 0.50 m below the open surface of the water and has a surface area of 8.3×10^{-3} m^2, what must the mass of the piston be to keep the system in mechanical equilibrium? For simplicity, ignore the mass of the partition.

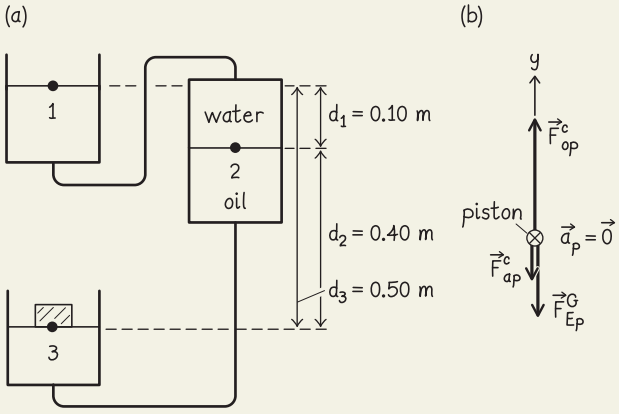
Figure 18.47 Example 18.7.



❶ **GETTING STARTED** I begin by making a sketch of the arrangement and identifying points of interest (**Figure 18.48***a*). Point 1 is at the water surface in the top tank, and the pressure here is atmospheric pressure. Point 2 is on the partition in the right tank separating the water from the oil. In mechanical equilibrium, the partition is not accelerating, and so, because I ignore the partition mass, the pressure must be the same on the two sides of the partition. Point 3 is at the position in the bottom tank where the oil and piston meet.

I also draw a free-body diagram for the piston, which also has zero acceleration in mechanical equilibrium

Figure 18.48



(Figure 18.48*b*). The piston is subject to an upward force $\vec{F}^c_{op}$ exerted by the oil, a downward force $\vec{F}^c_{ap}$ exerted by the air above the piston, and a downward gravitational force $\vec{F}^G_{Ep}$.

❷ **DEVISE PLAN** When the system is in mechanical equilibrium, the piston is stationary. Therefore the vector sum of the forces exerted on it must be zero. From my free-body diagram, I see that the magnitude of the upward force exerted by the oil on the piston is equal in magnitude to the sum of the downward gravitational force exerted on the piston and the downward force exerted by the air on the piston: $F^c_{op} = F^G_{Ep} + F^c_{ap} = m_p g + F^c_{ap}$.

The magnitudes F^c_{ap} and F^c_{op} are given by a rearranged form of Eq. 18.1: $F^c_{ap} = P_{air}A_{piston}$, and $F^c_{op} = P_3 A_{piston}$. The problem therefore reduces to determining the pressure P_3 at point 3, where the oil and piston meet. That point is 0.40 m below the partition in the tank on the right, where the pressure is P_2. Because points 2 and 3 are in a connected liquid (the oil), I can use Eq. 18.7 to relate P_2 and P_3. Points 1 and 2 are also in a connected liquid (the water), so I can use Eq. 18.7 to relate P_2 and $P_1 = P_{atm}$.

(Continued)

3 EXECUTE PLAN Because the vector sum of the forces exerted on the piston is zero, I have $F^c_{op} - F^c_{ap} - F^G_{Ep} = P_3 A_{piston} - P_{air} A_{piston} - m_{piston} g = 0$, or

$$m_{piston} = \frac{(P_3 - P_{air})A_{piston}}{g}. \qquad (1)$$

To obtain P_3, I first apply Eq. 18.7 to points 1 and 2, remembering that the lower point always goes on the left side of the equation: $P_2 = P_1 + \rho_{water} g d_1 = P_{air} + \rho_{water} g d_1$. Then I apply Eq. 18.7 to points 2 and 3: $P_3 = P_2 + \rho_{oil} g d_2 = P_{air} + \rho_{water} g d_1 + \rho_{oil} g d_2$. Substituting this expression into Eq. 1, I get

$$m_{piston} = \frac{(P_{air} + \rho_{water} g d_1 + \rho_{oil} g d_2 - P_{air})A_{piston}}{g}$$

$$= (\rho_{water} d_1 + \rho_{oil} d_2)A_{piston}$$

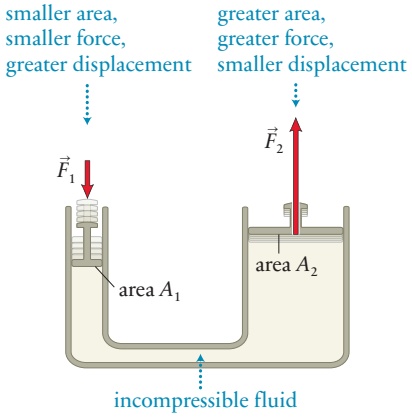
$$= [(1.0 \times 10^3 \text{ kg/m}^3)(0.10 \text{ m}) + (0.92 \times 10^3 \text{ kg/m}^3)(0.40 \text{ m})]$$

$$\times (8.3 \times 10^{-3} \text{ m}^2) = 3.9 \text{ kg}. \checkmark$$

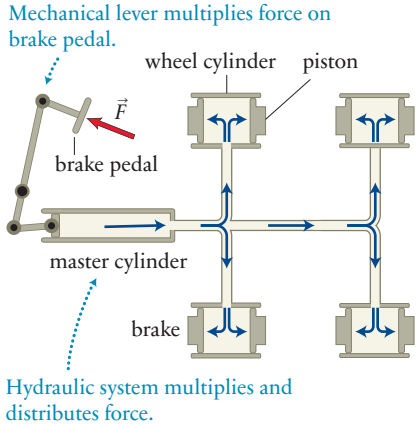
4 EVALUATE RESULT The atmospheric pressure term drops out because it affects both the water surface and the piston. Points 1 and 3 are separated by 0.5 m. Because I know that the pressure difference due to 10 m of water is equal to atmospheric pressure at sea level, or about 100 kPa (see Checkpoint 18.4; these are handy numbers to remember), I know that 0.5 m of water should correspond to a pressure difference of about 5 kPa. The mass density of oil is a bit smaller than that of water, but I can ignore the difference. The surface area of the piston is about 10^{-2} m^2, and so the force giving rise to the 5-kPa pressure difference is about 50 N, which is equal to the gravitational force exerted on a mass of 5 kg, which is close to the answer I obtained.

Figure 18.49 Hydraulic systems make it possible to lift heavy loads or exert large forces.

(*a*) Basic hydraulic system



(*b*) Schematic model of hydraulic brake system



18.16 Oil and water don't mix, and the mass density of oil is smaller than that of water. Suppose water is poured into a U-shaped tube that is open at both ends until the water surface is halfway up each leg of the tube, and then some oil is poured on top of the water in the right leg. Once the system comes to equilibrium, are the top of the oil column in the right leg and the top of the water column in the left leg at the same height? If not, which is higher?

The principles we have discussed in this section allow us to understand the working of *hydraulic systems*, which are used for lifting heavy loads or exerting large forces. The basic idea behind a hydraulic system is simple: A liquid, usually oil, is used to transmit a force applied at one point to another point. **Figure 18.49*a*** shows a schematic hydraulic system: two pistons fitted into two connected, oil-filled cylinders. Because the oil is incompressible, a force exerted on one piston increases the pressure everywhere in the oil, and so the force is transmitted to the other piston.

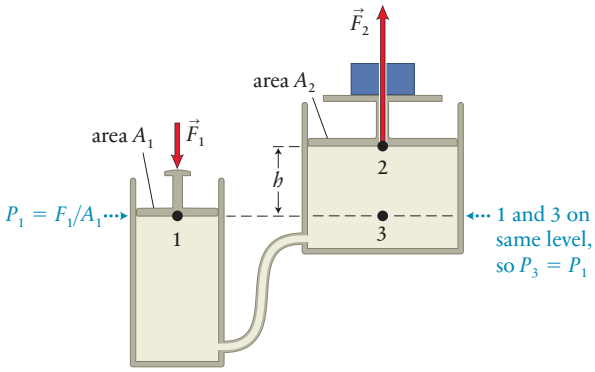
The tube that connects the two cylinders can be any length or shape, which means it can snake through tight spaces and can also fork, allowing an applied force to be distributed. In the brake system of a car, for example, the force your foot exerts on the brake pedal is transmitted to the piston in an oil-filled master cylinder. The pressure increase in the oil is transmitted through steel tubes to the four brakes in the wheels, where the oil exerts a force on the brake pads, pushing them against a disk attached to the wheel, which slows down the car (Figure 18.49*b*).

Another great advantage of hydraulic systems is that they permit us to "multiply" a force by varying the size of the pistons. In **Figure 18.50**, a force of magnitude F_1 is exerted on a small piston of area A_1, causing a gauge pressure $P_1 = F_1/A_1$ in the liquid at the small piston. Because $P_3 = P_2 + \rho g h$, the gauge pressure P_2 at the large piston is then

$$P_2 = P_3 - \rho g h = P_1 - \rho g h, \qquad (18.19)$$

QUANTITATIVE TOOLS

Figure 18.50 Hydraulic system used to "multiply" a force.



where h is the vertical distance between the pistons. When the liquid in a hydraulic system is oil, we can ignore the term ρgh because the mass density of oil is not very great and h is generally small, so that $P_1 \approx P_2$. The gauge pressure P_2 at point 2 causes the liquid to exert a force of magnitude F_2 on the piston. If the area of that piston is A_2, we have $P_2 = F_2/A_2$, so that

$$\frac{F_1}{A_1} = \frac{F_2}{A_2} \tag{18.20}$$

or

$$F_2 = \frac{A_2}{A_1} F_1. \tag{18.21}$$

Because $A_2 > A_1$, we have $F_2 > F_1$. Therefore a small force applied to the small piston causes the liquid to exert a large force on the large piston (even though the *pressure* is the same at both points). The factor A_2/A_1 is what permits you to exert large forces on the brakes of a car even though the force your foot exerts on the brake pedal is relatively small.

PROCEDURE: Working with pressure in liquids at rest

The branch of physics that deals with pressure in a liquid at rest is called *hydrostatics*. The pressure in a liquid at rest is determined by the force of gravity and by what happens at the boundary of the liquid. To determine the pressure in such liquids:

1. Begin by making a sketch showing all the **boundaries** and identifying all the factors that affect pressure: pistons, gases at surfaces open to the atmosphere, and so on. Note the known vertical heights of liquid surfaces, the areas of these surfaces, the surface areas of pistons, and the liquid mass densities.
2. Determine the **pressure at each surface,** The pressure at a liquid surface open to the air is equal to atmospheric pressure P_{atm}. The pressure at a liquid surface bordering a vacuum is zero: $P = 0$. The pressure at a liquid surface open to a gas other than the atmosphere is equal to the pressure in the gas: $P = P_{\text{gas}}$. The pressure at a liquid surface that is in contact with a solid, such as a piston, is $P = F^{\text{c}}_{\text{sl}}/A$, where F^{c}_{sl} is the magnitude of the force exerted by the solid on the liquid and A is the area over which that force is exerted.
3. Use **horizontal planes,** The pressure is the same at all points on a horizontal plane in a connected liquid. The pressure difference between two horizontal planes 1 and 2 is given by $P_1 = P_2 + \rho gd$ (Eq. 18.7), where d is the vertical distance between the horizontal planes and 1 is below 2.

QUANTITATIVE TOOLS

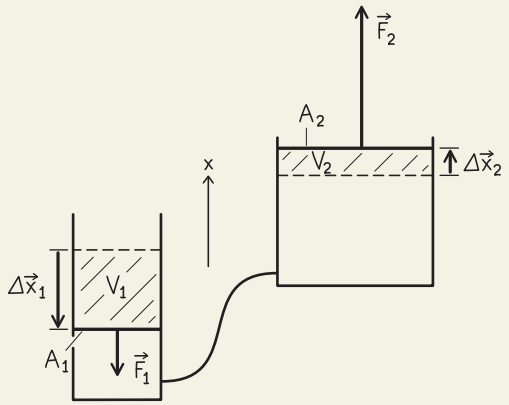
EXAMPLE 18.8 Liquid work

Suppose a load is placed on the large piston in Figure 18.50 and we wish to raise that load by a certain distance. How does the work done on the small piston compare with the work required to raise the load? Ignore the masses of the pistons.

1 GETTING STARTED I begin by making a sketch (**Figure 18.51**). To keep my subscripts simple, I call the small piston 1 and the large piston 2. I denote the displacements of the two pistons by $\Delta\vec{x}_1$ and $\Delta\vec{x}_2$, and the forces exerted on them by $\vec{F}_1$ and $\vec{F}_2$.

2 DEVISE PLAN The magnitudes F_1 and F_2 are related by Eq. 18.21, $F_2 = (A_2/A_1)F_1$. Each force does an amount of work given by the scalar product of the force and the force displacement, which is equal to the displacement of the piston on which the force is exerted. Both forces do positive work because the direction of each force is the same as the direction of the force displacement. Because the liquid in the cylinders is incompressible, I know that the volume V_1 of liquid pushed out of the small cylinder must be equal to the volume V_2 of liquid pushed into the large cylinder: $V_1 = V_2$. This expression will help me relate the piston displacements to the piston surface areas.

Figure 18.51



3 EXECUTE PLAN Raising the load a distance $d = \Delta x_2$ requires work $W_2 = \vec{F}_2 \cdot \Delta\vec{x}_2 = F_2 d$. Using Eq. 18.21, I can write for the work done on the small piston

$$W_1 = \vec{F}_1 \cdot \Delta\vec{x}_1 = F_1 \Delta x_1 = F_2 \frac{A_1}{A_2} \Delta x_1. \qquad (1)$$

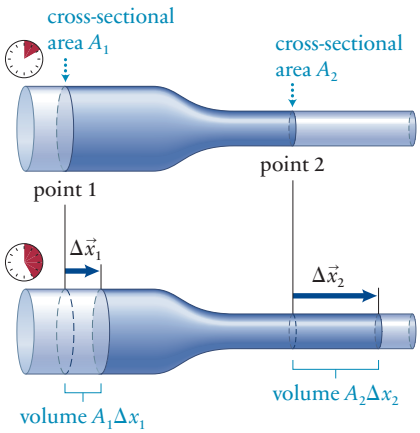
The volume of liquid pushed out of the small cylinder is $V_1 = A_1 \Delta x_1$, and that pushed into the large cylinder is $V_2 = A_2 \Delta x_2 = A_2 d$. Because these two volumes are equal, I have $A_1 \Delta x_1 = A_2 d$, or $\Delta x_1 = (A_2/A_1)d$. Substituting this expression into Eq. 1, I get $W_1 = F_2 d$, telling me that the same work is done on both pistons. ✔

4 EVALUATE RESULT My answer tells me that no energy is gained by using the hydraulic lift, which makes sense. The force exerted on the small piston is increased by the factor A_2/A_1 when it is transmitted to the large piston, but the displacement decreases by the inverse factor A_1/A_2, and so the work done is the same.

18.17 In the derivation of Eq. 18.21, we ignored the effect of atmospheric pressure on the two pistons. How is the result we obtained affected by atmospheric pressure?

18.7 Bernoulli's equation

Figure 18.52 A fluid flows through a tapered tube.



Next we turn our attention to fluids in motion. Consider the laminar flow of a nonviscous fluid through the tapered tube in **Figure 18.52**. In a time interval Δt, a thin "slice" of fluid in the wide part of the tube undergoes a displacement Δx_1. The volume of fluid that enters the tube at point 1 is thus $A_1|\Delta x_1|$, where A_1 is the cross-sectional area of the wide part of the tube. If the mass density of the fluid at point 1 is ρ_1, the amount of mass entering the tube at point 1 is $m_1 = \rho_1 A_1 |\Delta x_1|$. Likewise, if a thin slice of fluid in the narrow part of the tube has a displacement Δx_2, and the cross-sectional area of the narrow part is A_2, then the mass of the fluid emerging at point 2 is $\rho_2 A_2 |\Delta x_2|$. If the flow is laminar, the flow speed at any location doesn't change with time. This means that no mass can accumulate in the tube, and so the mass of the fluid emerging from the narrow part at point 2 during a time interval Δt must be the same as the mass entering at point 1:

$$\rho_1 A_1 |\Delta x_1| = \rho_2 A_2 |\Delta x_2| \quad \text{(laminar flow of nonviscous fluid)}. \qquad (18.22)$$

QUANTITATIVE TOOLS

If the flow speed of the fluid is v_1 at point 1 and v_2 at point 2, we have $|\Delta x_1| = v_1 \Delta t$ and $|\Delta x_2| = v_2 \Delta t$. Substituting these expressions into Eq. 18.22 and dividing both sides by Δt, we obtain

$$\rho_1 A_1 v_1 = \rho_2 A_2 v_2 \quad \text{(laminar flow of nonviscous fluid).} \tag{18.23}$$

Equation 18.23 is called the **continuity equation.** This equation quantitatively expresses the idea that the mass of the fluid emerging from the narrow part at point 2 during a time interval Δt is the same as the mass entering at point 1.

For an incompressible fluid, the mass density is the same in all parts of the tube and so

$$A_1 v_1 = A_2 v_2 \quad \text{(laminar flow of nonviscous, incompressible fluid).} \tag{18.24}$$

This equation expresses quantitatively what we concluded in Section 18.3: A fluid flows faster in a narrow part of a tube than in a wide part.

We can write Eq. 18.24 in another form by introducing a quantity Q, called the **volume flow rate,** the rate at which a volume of fluid crosses a section of a tube. If a volume V crosses a section in a time interval Δt, the volume flow rate is

$$Q \equiv \frac{V}{\Delta t}. \tag{18.25}$$

The SI unit of volume flow rate is meters cubed per second (m^3/s). For a nonviscous fluid flowing at speed v through a pipe of cross-sectional area A, we have $V = A|\Delta x| = Av\Delta t$, and so the volume flow rate is

$$Q = Av. \tag{18.26}$$

Equation 18.24 thus tells us that the volume flow rate is the same at all locations in an incompressible fluid: $Q_1 = Q_2$.

We also concluded in Section 18.3 that in order for a fluid to accelerate in the region where a tube narrows, the pressure must be greater in the wide part. Let us therefore now relate pressure to flow speed. Consider the tapered tube in **Figure 18.53**. At point 1 the tube's cross-sectional area is A_1, the fluid's speed is v_1, the pressure is P_1, and the vertical height above some arbitrarily chosen origin is y_1. The corresponding quantities at point 2 are A_2, v_2, P_2, and y_2. Let us consider the system made up of Earth and the fluid initially between points 1 and 2. How does the energy of this system change?

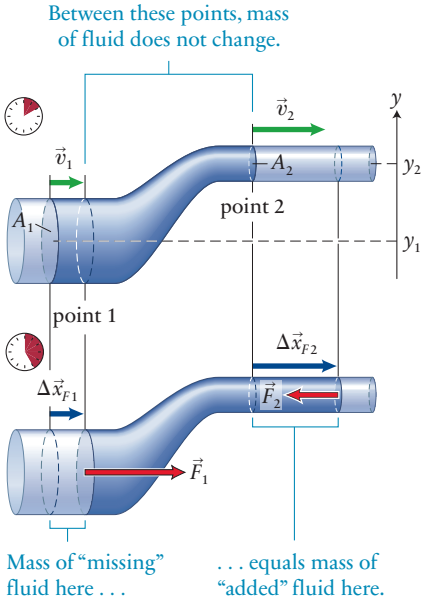
To answer this question, let us apply the energy law (Eq. 9.1, $\Delta E = W$) to this system over a time interval Δt. As Figure 18.53 shows, during the time interval Δt the system has shifted to the right. The difference between the initial and final states of the system is a "missing" volume of fluid on the left and an "added" volume of fluid on the right. The continuity equation requires the masses of these two volumes to be the same: $m_1 = m_2 = m$. Because the mass of the fluid between points 1 and 2 doesn't change, the only forms of energy we need to consider are the kinetic and potential energies associated with m_1 and m_2. The change in kinetic energy is

$$\Delta K = \tfrac{1}{2} m_2 v_2^2 - \tfrac{1}{2} m_1 v_1^2 = \tfrac{1}{2} m (v_2^2 - v_1^2), \tag{18.27}$$

and the change in gravitational potential energy is

$$\Delta U^G = m_2 g y_2 - m_1 g y_1 = mg(y_2 - y_1). \tag{18.28}$$

Figure 18.53 A volume of fluid flows through a tapered tube that changes altitude.



Substituting Eqs. 18.27 and 18.28 into the left side of the energy law yields

$$\Delta E = \Delta K + \Delta U^G = \tfrac{1}{2}m(v_2^2 - v_1^2) + mg(y_2 - y_1). \qquad (18.29)$$

Our system is not closed because the fluid that is in contact with the left and right termini of the system exerts forces $\vec{F}_1$ and $\vec{F}_2$ on the system and the points of application of these forces undergo nonzero force displacements $\Delta\vec{x}_{F1}$ and $\Delta\vec{x}_{F2}$,* respectively. So let us now evaluate the work W done on the system. We take the time interval Δt to be small enough so that we can consider the forces to be constant. Because $\vec{F}_1$ and the force displacement of $\vec{F}_1$ are in the same direction, the work done on the system at point 1 is positive:

$$W_1 = F_{1x}\Delta x_{F1} = +P_1A_1v_1\Delta t. \qquad (18.30)$$

At point 2, negative work is done on the system because the force and force displacement point in opposite directions:

$$W_2 = F_{2x}\Delta x_{F2} = -P_2A_2v_2\Delta t, \qquad (18.31)$$

and so the work done on the system is

$$W = P_1A_1v_1\Delta t - P_2A_2v_2\Delta t. \qquad (18.32)$$

Using Eq. 18.24, we can write $A_1v_1\Delta t = A_2v_2\Delta t$. The quantity $A_1v_1\Delta t$ is the volume of the mass of fluid m that flows into (or out of) the section of tube we are considering. From the definition of mass density, we know that $V = m/\rho$, which means that

$$\frac{m}{\rho} = A_1v_1\Delta t. \qquad (18.33)$$

Substituting this expression into Eq. 18.32, we get

$$W = \frac{m}{\rho}(P_1 - P_2). \qquad (18.34)$$

Because the work done on the system must be equal to the change in the system's energy, we have, from Eqs. 18.29 and 18.34:

$$\tfrac{1}{2}m(v_2^2 - v_1^2) + mg(y_2 - y_1) = \frac{m}{\rho}(P_1 - P_2) \qquad (18.35)$$

or, dividing both sides by m, rearranging terms, and multiplying through by ρ,

$$P_1 + \rho g y_1 + \tfrac{1}{2}\rho v_1^2 = P_2 + \rho g y_2 + \tfrac{1}{2}\rho v_2^2 \quad \text{(laminar flow of incompressible, nonviscous fluid).} \qquad (18.36)$$

Together with the continuity equation (Eq. 18.23), Eq. 18.36, known as **Bernoulli's equation,** forms the foundation of the hydrodynamics of

*The subscript F on the displacement means we are considering the displacement of the point of application of the force (see Section 9.1).

QUANTITATIVE TOOLS

incompressible fluids. Strictly speaking, Bernoulli's equation applies only to liquids because we used Eq. 18.24, which is valid only for incompressible fluids. However, for the flow of a gas at modest velocities (and for height differences $y_2 - y_1$ that are not too great), the pressure variations have little effect on the mass density and Bernoulli's equation is still a reasonable approximation.

Because the points 1 and 2 are arbitrary, Eq. 18.36 tells us that the quantity $P + \rho g y + \frac{1}{2}\rho v^2$ is the same at different locations in a flow stream. If the fluid is stationary, $v = 0$ everywhere, and Eq. 18.36 takes on the same form as Eq. 18.6. If points 1 and 2 are at the same height, then $y_1 = y_2$, and Eq. 18.36 becomes

$$P_1 + \tfrac{1}{2}\rho v_1^2 = P_2 + \tfrac{1}{2}\rho v_2^2 \quad \begin{array}{l}\text{(horizontal laminar flow}\\ \text{of incompressible,}\\ \text{nonviscous fluid).}\end{array} \qquad (18.37)$$

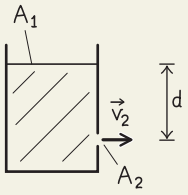
This equation confirms what we concluded in Section 18.3: The pressure P is greater where the flow speed v is smaller, and vice versa.

EXAMPLE 18.9 Leaky bucket

Water leaks out of a small hole in the side of a bucket. The hole is a distance d below the surface of the water, and the cross section of the hole is much smaller than the diameter of the bucket. At what speed does the water emerge from the hole?

❶ **GETTING STARTED** I begin by making a sketch of the bucket (**Figure 18.54**), defining my point 1 at the water surface and my point 2 at the hole. The bucket can be regarded as a tube open at both ends, with the top opening, cross-sectional area A_1, as one end and the hole, cross-sectional area A_2, as the other end.

Figure 18.54



❷ **DEVISE PLAN** Because $A_1 \gg A_2$, the continuity equation (Eq. 18.23, $\rho_1 A_1 v_1 = \rho_2 A_2 v_2$) tells me that the speed v_1 at which the surface of the water moves downward is much smaller than the speed v_2 at which the water emerges from the hole, $v_1 \ll v_2$. For all practical purposes, therefore, I can assume $v_1 \approx 0$. I also know that the pressure both at the water surface and at the hole is equal to atmospheric pressure. I can then use Bernoulli's equation (Eq. 18.36) to determine v_2.

❸ **EXECUTE PLAN** Because $P_1 = P_2 = P_{\text{atm}}$, Bernoulli's equation reduces to $\rho g y_1 + \frac{1}{2}\rho v_1^2 = \rho g y_2 + \frac{1}{2}\rho v_2^2$. Dividing through by $\frac{1}{2}\rho$, setting $v_1 = 0$, and bringing the terms containing y to the left side, I get $v_2^2 = 2g(y_1 - y_2) = 2gd$, and so $v_2 = \sqrt{2gd}$. ✔

❹ **EVALUATE RESULT** My answer is equal to the speed acquired over a vertical distance d by a freely falling object starting from rest. Because the Earth-object system is closed, $\Delta K + \Delta U^G = \frac{1}{2}mv^2 - mgd = 0$, and so $v^2 = 2gd$. My result makes sense: For each water drop that emerges from the hole, the water at the surface is reduced by an equal amount and the kinetic energy of the emerging drop must be equal to the decrease in potential energy.



18.18 Suppose that the diameter of the hole in Example 18.9 is 10.0 mm, the bucket diameter is 0.50 m, and the distance from the hole to the water surface is 0.30 m. At what speed does the water emerge from the hole?

18.8 Viscosity and surface tension

So far in our discussion of fluid flow we have ignored energy dissipation. As pointed out in Section 18.3, viscosity is a measure of a fluid's resistance to shear stress. Dissipation of energy occurs in both liquids and gases, and so both have a nonzero viscosity. Consider, for example, two flat plates

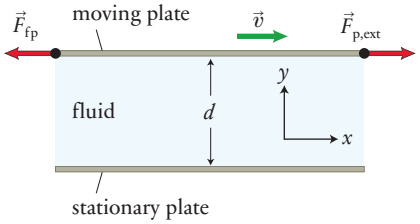


Figure 18.55 To measure viscosity, a fluid is placed between two parallel plates. The bottom plate is held fixed while the top plate is translated. The force required to keep the top plate moving at constant speed gives the viscosity of the fluid.

separated by a fluid (**Figure 18.55**). The lower plate is held fixed while the upper plate is translated in the plane of the plate. Because of friction between adjacent fluid layers, a force $\vec{F}_{p,ext}$ is required to move the upper plate at constant speed v. Experiments show that the magnitude of this force is proportional to the surface area A of the upper plate and to the speed v, and inversely proportional to the distance d between the plates. Because the force is directed in the direction opposite that of the velocity of the upper plate, we can write for a plate moving in the positive x direction:

$$F_{fpx} \equiv -\eta A \frac{v_x}{d}. \tag{18.38}$$

The proportionality constant η (Greek lowercase eta) is the **viscosity** of the fluid. The SI unit for viscosity is the pascal-second, with $1\ \text{Pa}\cdot\text{s} = 1\ \text{kg}/(\text{m}\cdot\text{s})$. The greater the viscosity of a fluid, the greater the force needed to move the plate at constant speed. Because the work done by the plate on the fluid ends up as thermal energy in the fluid, a greater viscosity means more energy dissipation in the fluid as fluid layers rub past one another.

The ratio v_x/d in Eq. 18.38 can also be written as dv_x/dy because both v_x/d and dv_x/dy express the change in the x component of the velocity divided by the distance over which the change occurs. The derivative dv_x/dy, called the *velocity gradient*, is a measure of the rate at which one fluid layer moves relative to an adjacent layer. The advantage of expressing Eq. 18.38 in terms of the velocity gradient is that the resulting equation can be applied more generally:

$$F_{fpx} = -\eta A \frac{dv_x}{dy}. \tag{18.39}$$

Because viscous forces are generally small, fluids are often used as lubricants to reduce friction between solids. **Table 18.1** lists some typical viscosities. The viscosity of gases is typically several orders of magnitude smaller than that of liquids. The viscosity of liquids typically decreases as temperature increases. Honey, for example, becomes more runny as it is heated. In contrast, gases typically get more viscous at higher temperatures.

Table 18.1 Viscosity of fluids

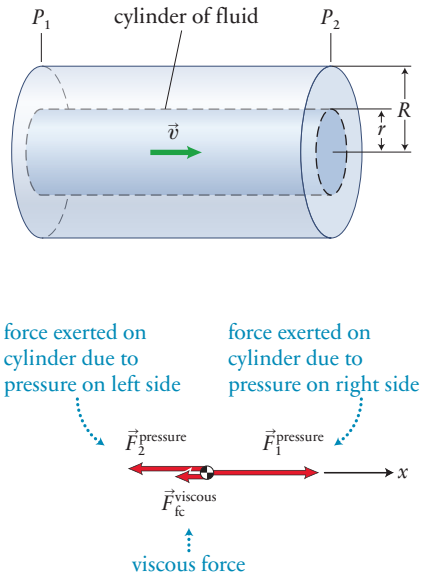
Fluid	Viscosity (Pa · s)	Temperature (°C)
Air	1.71×10^{-5}	0
	1.81×10^{-5}	20
	2.18×10^{-5}	100
Blood	3.0×10^{-3}	20
	2.0×10^{-3}	37
Castor oil	5.3	0
	0.986	20
	0.017	100
Water	1.79×10^{-3}	0
	1.00×10^{-3}	20
	0.282×10^{-3}	100

Many instances of fluid flow, such as the velocity distribution of water through pipes or the flow of blood through arteries, involve laminar flow through cylindrical tubes. Let us therefore consider the flow of a fluid through a horizontal cylindrical pipe of length ℓ and radius R. Because the boundary layer of the fluid sticks to the interior wall of the pipe, this layer is at rest. The layer of fluid adjacent to the boundary layer moves slowly in the direction of the flow, and adjacent layers move at increasing speeds, with the fluid at the center of the pipe having the maximum speed (see Figure 18.19*b*).

Consider a cylinder of fluid of radius r flowing in the positive x direction through a tube of radius $R > r$. The dashed outline in **Figure 18.56** shows such a cylinder of fluid. Because of friction between the fluid in the cylinder and the more slowly moving layer of fluid adjacent to the cylinder boundary, a viscous force $\vec{F}_{fc}^{\,viscous}$ is exerted on the cylinder in the negative x direction. To keep the cylinder moving at constant velocity through the tube, a force of equal magnitude must be exerted on it in the positive x direction. This force requires a pressure difference ΔP between the ends of the cylinder, and so we see that for a fluid with nonzero viscosity, pressure must decrease as the fluid moves along the tube. For motion at constant speed, we have

$$\Sigma \vec{F}_c = \vec{F}_{fc}^{\,pressure} + \vec{F}_{fc}^{\,viscous} = 0. \tag{18.40}$$

Figure 18.56 In mechanical equilibrium, the force due to the pressure on the two ends of the fluid in the cylinder defined by the dashed lines must balance the viscous force opposing the motion of the cylinder.



If the pressure at the left end of the cylinder is P_1, the fluid to the left of this end exerts a force of magnitude $F_1^{pressure} = P_1(\pi r^2)$ in the positive x direction, where πr^2 is the cross-sectional area of the cylinder. If the pressure at the right end of the cylinder is P_2, the fluid to the right of this end exerts a force of magnitude $F_2^{pressure} = P_2(\pi r^2)$ in the negative x direction. The vector sum of these two forces is $(P_1 - P_2)(\pi r^2)$. The viscous force $F^{viscous}$ on the cylinder is given by Eq. 18.39, with A equal to the surface area $2\pi r\ell$ of the cylinder of fluid and $dv_x/dy = -dv_x/dr$.* Substituting this information into Eq. 18.40, we get

$$(P_1 - P_2)(\pi r^2) + \eta(2\pi r\ell)\frac{dv_x}{dr} = 0 \tag{18.41}$$

or, rearranging terms,

$$-dv_x = \frac{(P_1 - P_2)(\pi r^2)}{\eta(2\pi r\ell)}dr = \frac{(P_1 - P_2)}{2\eta\ell}rdr. \tag{18.42}$$

If we integrate the right side from r to R, the x component of velocity goes from v_x (the value at r) to 0 (the value at R, where the fluid is at rest because its boundary layer sticks to the tube wall):

$$-\int_{v_x}^{0} dv_x = \frac{P_1 - P_2}{2\eta\ell}\int_r^R rdr = \frac{P_1 - P_2}{2\eta\ell}\left[\tfrac{1}{2}r^2\right]_r^R \tag{18.43}$$

or

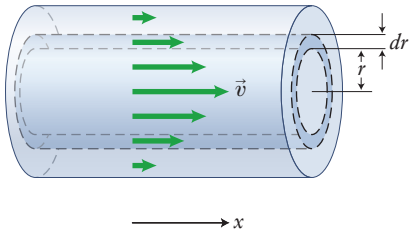
$$v_x = \frac{P_1 - P_2}{4\eta\ell}(R^2 - r^2) \quad \text{(laminar flow in cylindrical tube).} \tag{18.44}$$

*In Figure 18.19*b* v_x decreases with increasing r, and so dv_x/dr is negative. In Figure 18.55 v_x increases with increasing y, and so dv_x/dy is positive.

This equation shows that the velocity varies with the square of r, and so its curve has a parabolic shape (see Figure 18.19*b*) with a maximum at the center of the tube, where $r = 0$.

Because the fluid velocity varies with r, the volume flow rate is no longer simply Av. To calculate the volume flow rate, we consider a thin cylindrical shell of thickness dr (**Figure 18.57**). The shell has a cross-sectional area $dA = 2\pi r dr$ and contributes to the overall volume flow rate an amount

Figure 18.57 Each cylindrical shell (dashed lines) contributes an amount dQ to the volume flow rate through the tube.



$$dQ = v_x dA = v_x(2\pi r dr) = \frac{P_1 - P_2}{4\eta\ell}(R^2 - r^2)(2\pi r dr)$$

$$= \frac{\pi(P_1 - P_2)}{2\eta\ell}(R^2 r - r^3)dr, \qquad (18.45)$$

where we have substituted Eq. 18.44 for v_x. To obtain the volume flow rate Q through the tube, we must integrate this expression from $r = 0$ to $r = R$. Integrating the factor $(R^2 r - r^3)$ in Eq. 18.45 yields

$$\int_0^R (R^2 r - r^3)dr = \left[\tfrac{1}{2}R^2 r^2 - \tfrac{1}{4}r^4\right]_0^R = \tfrac{1}{2}R^4 - \tfrac{1}{4}R^4 = \tfrac{1}{4}R^4, \qquad (18.46)$$

and so the volume flow rate through a cylindrical tube is

$$Q = \frac{\pi R^4}{8\eta\ell}(P_1 - P_2) \quad \text{(laminar flow in cylindrical tube).} \qquad (18.47)$$

Equation 18.47, called **Poiseuille's law,** shows that a high viscosity leads to a low-volume flow rate, which is reasonable. The strong dependence on the tube radius R means that a small adjustment in radius can lead to a large increase in volume flow rate. For example, increasing R by just 20% increases the volume flow rate by a factor of $(1.2)^4 = 2.1$.

EXAMPLE 18.10 Blood flow

Blood flows through an artery that has a radius of 3.0 mm. What is the maximum flow speed if the volume flow rate through the artery is 2.0×10^{-6} m^3/s?

❶ **GETTING STARTED** I assume that the artery is cylindrical, which means I can use the results from this section.

❷ **DEVISE PLAN** I can obtain the maximum flow speed by setting $r = 0$ in Eq. 18.44. This equation contains the unknown factor $(P_1 - P_2)/(4\eta\ell)$. The expression for the volume flow rate, Eq. 18.47, however, contains the same factor, which gives me two equations with two unknowns—the maximum flow speed v and the factor $(P_1 - P_2)/(4\eta\ell)$—so I can solve for both unknowns.

❸ **EXECUTE PLAN** Setting $r = 0$ in Eq. 18.44 gives $v_{max} = (P_1 - P_2)R^2/(4\eta\ell)$. Next, I rewrite Eq. 18.47 in the form $(P_1 - P_2)/(4\eta\ell) = 2Q/(\pi R^4)$, which I substitute into my expression for v_{max}:

$$v_{max} = \frac{(P_1 - P_2)R^2}{4\eta\ell} = \frac{2Q}{\pi R^4}R^2 = \frac{2Q}{\pi R^2}$$

$$= \frac{2(2.0 \times 10^{-6}\ \text{m}^3/\text{s})}{\pi(3.0 \times 10^{-3}\ \text{m})^2} = 0.14\ \text{m/s}. \checkmark$$

❹ **EVALUATE RESULT** I can rewrite the equality $v_{max} = (2Q)/(\pi R^2)$ from my final equation in the form $Q = \frac{1}{2}\pi R^2 v_{max} = \frac{1}{2}Av_{max}$, where A is the cross-sectional area of the artery. This equation is very similar to Eq. 18.26, the definition of volume flow rate. Indeed, $\frac{1}{2}v_{max}$ is the average flow speed. The value I obtain also appears to be reasonable—a given volume of blood travels about 1 m in 7 s.

Having dealt with forces in the bulk of a fluid, let us now turn to forces at liquid surfaces. As we saw in Section 18.4, these forces cause surface tension at the surface and are responsible for the shape of drops of liquid and capillary rise.

QUANTITATIVE TOOLS

Figure 18.58 shows a device commonly used to measure surface tension, a length of metal bent into a U shape and attached to a loosely fitted wire to form a rectangular loop. The wire fits so loosely that it slides off when the loop is held with the U inverted. When the loop is dipped into a soap solution, however, the soap film that forms inside the loop holds the wire stationary, as Figure 18.58 shows. An upward force must be exerted on the wire to hold it in place by balancing the downward force of gravity. The only likely candidate for this upward force is the one exerted by the film, so that $F_{\text{fw}} = F^G_{\text{Ew}}$. Experiments show that F_{fw} remains constant as we move the wire up or down. Because the thickness of the film changes as the wire is moved, the observation that the force magnitude F_{fw} does not change means that the film thickness is unimportant. Therefore we must attribute the upward force exerted on the wire to the front and back surfaces of the film. Experimentally we also find that the magnitude of this force is proportional to the length ℓ of the wire. The constant of proportionality

$$\gamma \equiv \frac{F_{\text{fw}}}{2\ell} \tag{18.48}$$

is called the **surface tension** of the liquid that makes up the film. The factor 2 in the denominator accounts for the two surfaces of the film.

The SI unit for surface tension is newtons per meter. **Table 18.2** gives the surface tension for a number of liquids in contact with air.

We can also understand surface tension in terms of the energy associated with the surface of a liquid. Moving the wire in Figure 18.58 causes a force displacement Δx_F, and so work $W = F_{\text{fw}x}\Delta x_F$ is done by the film on the wire. Because $\vec{F}_{\text{fw}}$ points in the positive x direction, we have $F_{\text{fw}x} = +2\ell\gamma$, and so $W = 2\ell\gamma\Delta x_F = -2\gamma\Delta A$, where ΔA is the change in the surface area A of the film. (When the wire moves, Δx_F and ΔA have opposite algebraic signs, so $\ell\Delta x_F = -\Delta A$.) As we move the wire down, we increase the surface area A of the film, and so $\Delta A > 0$ and the work done by the film on the wire is negative. Consequently the work done by the wire on the film is positive (the energy of the film increases). Where does the energy associated with this work go? As we increase the surface area, more liquid particles

Figure 18.58 The force exerted by a liquid film on a wire is due to surface tension in the film.

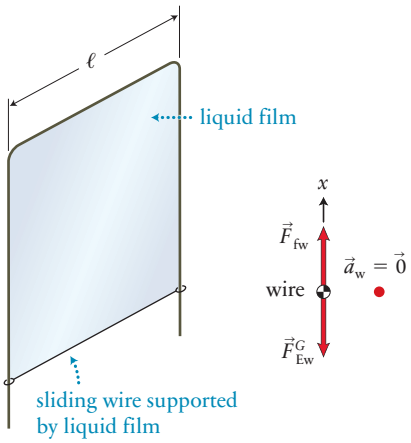
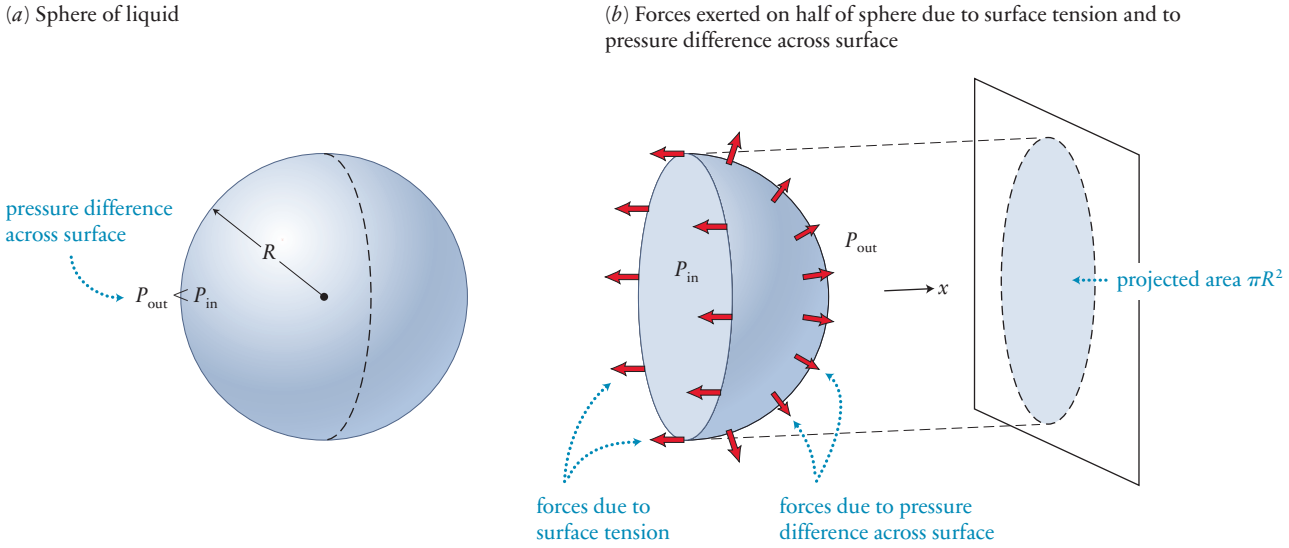


Table 18.2 Surface tensions of liquids in contact with air

Liquid	Surface tension (N/m)	Temperature (°C)
alcohol	2.23×10^{-2}	20
copper	1.100	1130
glycerin	6.31×10^{-2}	20
gold	1.000	1070
mercury	0.465	20
neon*	5.15×10^{-3}	−247
olive oil	3.20×10^{-2}	20
oxygen*	1.57×10^{-2}	−193
silver	0.800	970
water	7.57×10^{-2}	0
	7.28×10^{-2}	20
	5.89×10^{-2}	100

*Values given for liquid in contact with vapor.

Figure 18.59 Model for deriving the relationship between surface tension and pressure difference across a curved liquid surface. The surface tension forces are all directed perpendicular to the circular cross section, and the pressure-difference forces are all directed outward.



must be brought to the surface. For each particle brought to the surface, we reduce the number of bonds between particles (see Figure 18.25), and so the energy associated with the work we do in moving the wire downward goes into breaking bonds between neighboring particles in order to bring them to the surface. So, another way of understanding surface tension is to see it as the energy required to increase the surface area of a liquid. The greater the surface tension of a liquid, the more energy required to increase the surface area of a sample of the liquid by a certain amount.

As we saw in Section 18.4, surface tension is responsible for the pressure difference across a curved liquid surface. To derive the relationship between surface tension and pressure difference, consider the spherical body of liquid in **Figure 18.59*a***. The radius of the sphere is R, and the liquid has a surface tension γ. The pressure inside the liquid is P_{in}, and the pressure outside is P_{out}. Let's consider the forces exerted on one hemisphere, as illustrated in Figure 18.59*b*. Because the hemisphere remains at rest, the vector sum of the outward force due to the pressure difference across the surface and the inward force due to the surface tension must be zero:

$$\Sigma F_{\text{hemi } x} = F_x^{\text{pressure}} + F_x^{\text{surf tension}} = 0. \tag{18.49}$$

The force due to the pressure difference is perpendicular to the hemisphere surface. The components of the forces that are exerted on each segment of the surface but are not in the x direction cancel, and the x components of all these forces sum to the pressure difference $P_{in} - P_{out}$ times πR^2, which is the area of the circle projected on the screen in Figure 18.59*b*. The force due to the surface tension is exerted along the perimeter of the hemisphere, which has length $\ell = 2\pi R$. Therefore the magnitude of the force due to surface tension is $F_x^{\text{surf tension}} = 2\pi R\gamma$ (from Eq. 18.48, $\gamma = F/2\ell$; using this equation would give us $F = 2\ell\gamma = 2(2\pi R)\gamma = 4\pi R\gamma$, but we have to drop the factor 2 from Eq. 18.48 because here we are working with only one surface). Substituting these forces into Eq. 18.49 gives

$$(P_{in} - P_{out})\pi R^2 - 2\pi R\gamma = 0 \tag{18.50}$$

QUANTITATIVE TOOLS

or, rearranging terms,

$$P_{\text{in}} - P_{\text{out}} = \frac{2\gamma}{R} \quad \text{(spherical surface)}. \tag{18.51}$$

Equation 18.51 is a quantitative form of **Laplace's law** for a spherical surface.

For a cylindrical surface, it can be shown that Eq. 18.51 takes the form

$$P_{\text{in}} - P_{\text{out}} = \frac{\gamma}{R} \quad \text{(cylindrical surface)}. \tag{18.52}$$

EXAMPLE 18.11 Transmural pressure

The pressure difference across the wall of a blood vessel is called *transmural pressure*. Suppose an aorta of radius 11 mm has a transmural pressure of 12 kPa and a capillary of radius 4.0 μm has a transmural pressure of 5.0 kPa. Which blood vessel requires a greater tensile stress (tensile force per unit length) in its wall?

❶ **GETTING STARTED** I consider the aorta and the capillary to be elastic cylindrical tubes. Because the wall of a blood vessel keeps the blood confined within the tube, the tension in the wall plays the role of the surface tension γ of an unconfined liquid.

❷ **DEVISE PLAN** To determine γ for each vessel, I can use Eq. 18.52, which relates the pressure difference across the wall to the radius of the cylinder.

❸ **EXECUTE PLAN** I first rewrite Eq. 18.52 in the form $\gamma = (P_{\text{in}} - P_{\text{out}})R$. For the aorta,

$$\gamma = (12\ \text{kPa})\left(\frac{1000\ \text{Pa}}{1\ \text{kPa}}\right)(11\ \text{mm})\left(\frac{1\ \text{m}}{1000\ \text{mm}}\right)$$

$$= 1.3 \times 10^2\ \text{N/m}.$$

(You should be able to confirm that the units in this expression are indeed newtons per meter.) For the capillary,

$$\gamma = (5.0\ \text{kPa})\left(\frac{1000\ \text{Pa}}{1\ \text{kPa}}\right)(4.0\ \mu\text{m})\left(\frac{1\ \text{m}}{10^6\ \mu\text{m}}\right)$$

$$= 0.020\ \text{N/m}.$$

The tensile force per unit length in the aorta is much greater than that in the capillary. ✔

❹ **EVALUATE RESULT** The pressure differences across the aorta and capillary are of the same order of magnitude, and so less tension is required when the radius is small. This is the same situation as in Figure 18.27: The pressure difference is the same everywhere on the balloon surface, but the tension is less where the radius of curvature is smaller.



18.19 What is Laplace's law, Eq. 18.51, for a soap bubble?

Chapter Summary

Pressure due to a fluid (Sections 18.1, 18.5, 18.6)

Concepts A **fluid** is a substance that has no fixed shape and the ability to flow. A *nonviscous* fluid is one that has no resistance to deformation by shear stress.

Pascal's principle states that a pressure change applied to an enclosed liquid is transmitted undiminished to every part of the liquid and to the walls of the container.

Quantitative Tools If a fluid exerts a force $\vec{F}^{\,c}_{fs}$ perpendicular to a surface of area A, the **pressure** P on this area is a scalar defined by

$$P \equiv \frac{F^{c}_{fs}}{A}. \tag{18.1}$$

The SI unit of pressure is the **pascal** (Pa):

$$1\ \text{Pa} \equiv 1\ \text{N/m}^2. \tag{18.2}$$

The atmospheric pressure at sea level is 1.01325×10^5 Pa.

The pressure P at a depth d below the surface of a liquid at rest is independent of the shape of the container and is

$$P = P_{\text{surface}} + \rho g d, \tag{18.8}$$

where P_{surface} is the pressure at the surface of the liquid, g is the acceleration due to gravity, and ρ is the mass density of the liquid.

When pressure is measured, the actual pressure P is the sum of *gauge pressure* (pressure in excess of atmospheric pressure) and atmospheric pressure:

$$P = P_{\text{gauge}} + P_{\text{atm}}. \tag{18.16}$$

Buoyancy (Sections 18.2, 18.5)

Concepts If an object is fully or partially submerged in a fluid, the fluid exerts an upward **buoyant force** on the object. **Archimedes' principle** states that the buoyant force is equal in magnitude to the force of gravity exerted on the fluid displaced by the object. The volume of the displaced fluid is equal to the volume of the submerged part of the object.

Quantitative Tools The **buoyant force** $\vec{F}^{\,b}_{fo}$ that a fluid of mass density ρ_f exerts on an object that displaces a volume V_{disp} of fluid has magnitude

$$F^{b}_{fo} = \rho_f V_{\text{disp}} g. \tag{18.12}$$

Laminar flow of nonviscous fluids (Sections 18.3, 18.7)

Concepts In **laminar flow,** the velocity of the fluid at any given location is constant. Chaotically changing flow is called **turbulent flow.**

The mass density of an incompressible fluid is the same at all locations in the fluid.

Quantitative Tools The **continuity equation** for laminar flow of a nonviscous fluid through a tube is

$$\rho_1 A_1 v_1 = \rho_2 A_2 v_2, \tag{18.23}$$

where ρ, A, and v are the fluid mass density, tube cross-sectional area, and fluid speed, respectively, at any given location.

The **volume flow rate** Q (m^3/s) of a volume V of the fluid flowing past a given location in the tube in a time interval Δt is

$$Q \equiv \frac{V}{\Delta t}. \tag{18.25}$$

The volume flow rate can also be expressed as

$$Q = Av, \tag{18.26}$$

where A is the cross-sectional area of the tube and v is the fluid speed.

Bernoulli's equation describing laminar flow in an incompressible, nonviscous fluid is

$$P_1 + \rho g y_1 + \tfrac{1}{2}\rho v_1^2 = P_2 + \rho g y_2 + \tfrac{1}{2}\rho v_2^2, \tag{18.36}$$

where y is the vertical distance above an arbitrarily chosen reference location.

Viscosity and surface tension (Sections 18.4, 18.8)

Concepts **Viscosity** is a scalar measure of a fluid's resistance to shear deformation.

Surface tension is a scalar that gives the magnitude of the force per unit length exerted by a liquid surface on the perimeter of the surface.

Quantitative Tools The **viscosity** η (Pa · s) of a fluid confined between a lower fixed plate and an upper plate that is free to move is defined by

$$F_{\text{fp}x} \equiv -\eta A \frac{v_x}{d}, \tag{18.38}$$

where $F_{\text{fp}x}$ is the x component of the force $\vec{F}_{\text{fp}}$ exerted by the fluid on the upper plate when it moves in the plane of the plate with speed v_x, A is the plate surface area, and d is the vertical distance between the plates. More generally, we can write

$$F_{\text{fp}x} = -\eta A \frac{dv_x}{dy}, \tag{18.39}$$

where dv_x/dy is the rate at which one infinitesimally thin fluid layer moves past an adjacent layer.

Poiseuille's law: If P_1 and P_2 are the pressures at opposite faces of a cylindrical tube of radius R and length ℓ, the flow rate through the tube is

$$Q = \frac{\pi R^4}{8\eta\ell}(P_1 - P_2). \tag{18.47}$$

If a liquid exerts a force $\vec{F}_{\text{fw}}$ on a length ℓ of wire, the **surface tension** γ (N/m) of the liquid is

$$\gamma \equiv \frac{F_{\text{fw}}}{2\ell}. \tag{18.48}$$

Laplace's law states that the pressure difference $P_{\text{in}} - P_{\text{out}}$ across a curved liquid surface of radius of curvature R is

$$P_{\text{in}} - P_{\text{out}} = \begin{cases} \dfrac{2\gamma}{R} \text{ (spherical surface)} & (18.51) \\ \dfrac{\gamma}{R} \text{ (cylindrical surface).} & (18.52) \end{cases}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

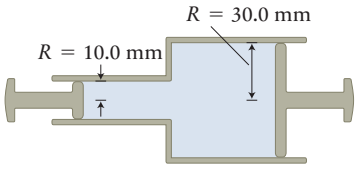
Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

Unless told otherwise, use 1000 kg/m^3 for the mass density of water.

18.1 Forces in a fluid

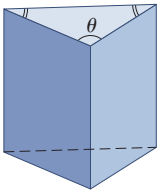
1. The liquid-filled cylinder in Figure P18.1 is clamped in place and has a piston on either end. The piston on the right is cylindrical and has a radius of 30.0 mm. The piston on the left is also cylindrical and has a radius of 10.0 mm. If the right cylinder is pushed leftward with a force of 30.0 N, what force must be exerted on the left cylinder to keep it from moving? •

Figure P18.1



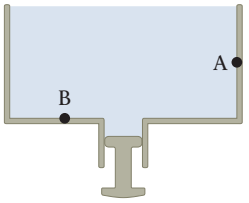
2. On Earth, a cylindrical container of water open to the atmosphere experiences atmospheric pressure at the top surface and a greater pressure at the bottom of the container. How do these pressures change when the container is being carried on the orbiting International Space Station? •
3. A box has six faces with three distinct areas $A < B < C$. Assume that A and B represent the four vertical surfaces, and the two horizontal surfaces have area C. When the box is submerged, rank the faces in order of (*a*) increasing water pressure and (*b*) increasing force exerted by the water. ••
4. You are designing a wedge-shaped container as illustrated in Figure P18.4. You must choose a value for the angle θ, which then determines the other two equal angles in this isosceles wedge. (*a*) When the container is filled with a fluid, what must the value of θ be if F/A, the ratio of the magnitude of the force exerted by the fluid on a side to the area of that side, is the same for all three sides? (*b*) What must θ be in order to have the magnitude of the force exerted by the fluid be the same on all three sides? ••

Figure P18.4



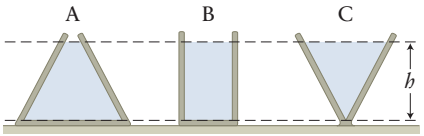
5. The open fixed container in Figure P18.5 is filled with a liquid and fitted with a piston at the bottom. Give the direction (if any) of the following: (*a*) the force $\vec{F}^{\,c}_{\ell w}$ exerted by the liquid on the wall at location A, (*b*) the pressure P in the liquid at A, (*c*) the force $\vec{F}^{\,c}_{w\ell}$ exerted by the wall on the liquid at location B, (*d*) the force $\vec{F}^{\,c}_{\ell p}$ exerted by the liquid on the piston if the piston is pushed, and (*e*) the force $\vec{F}^{\,c}_{\ell p}$ exerted by the liquid on the piston if the piston is pulled. (*f*) Which answers would change if the container were filled with a gas rather than a liquid? ••

Figure P18.5



6. You and your snowboard, with a combined mass of 70 kg, ride a half-pipe that has a radius of 5.0 m. As you pass the bottom, you create a pressure of 27 kN/m^2 on the half-pipe while moving at 8.2 m/s. If your snowboard is 1.3 m long, what is the average width of the part of your snowboard that is in contact with the half-pipe? ••
7. For the floor in any room, the *floor loading* is the maximum safe *average* pressure on the floor (the gravitational force exerted by Earth on all objects resting on the floor divided by the floor area), and the floor's *point loading* is the maximum safe *actual* pressure (the gravitational force exerted by Earth on all objects resting on the floor divided by the floor area covered by those objects). A particular room has a floor loading of 3 kN/m^2 and a point loading of 60 kN/m^2. What fraction of the floor area can be subjected to the point loading without exceeding the floor loading? ••
8. Three jars each have volume V and height h, but they have different shapes as shown in Figure P18.8. Water is to be added to each jar and the pressure is to be measured right at the bottom of each jar. Rank the jars in order of increasing pressure if the volume of water is (*a*) V and (*b*) $V/2$. ••

Figure P18.8



9. In a galaxy far, far away, a planet composed of an incompressible liquid of uniform mass density ρ has mass m_{planet} and radius R. Determine the pressure midway between the surface and the center of the planet. •••

18.2 Buoyancy

10. A cube measuring 100 mm on a side floats upright in water with 70.0 mm submerged. What is the mass of the cube? •
11. A scuba diving weight has a mass of 5.0 kg and is made from a material that has a mass density of 2000 kg/m^3. How much force is required to lift it from the bottom of a shallow pool? •
12. You have three bricks made of lead, stone, and plastic. The bricks have identical dimensions. When completely submerged (*a*) which brick feels the greatest buoyant force, and (*b*) which is most likely to float when released? •
13. A ship constructed of 2.50×10^6 kg of steel is roughly shaped like a box of length 50.0 m, width 20.0 m, and height 20.0 m. (*a*) What is the mass density of the ship? (*b*) Does the ship float? (*c*) If it floats, what is the maximum cargo mass it can carry without sinking? ••
14. A wooden block of height h is pushed from the air into a bucket of water that has a depth equal to $2h$. Make a graph showing how the buoyant force exerted on the block changes as its bottom moves from the surface to the bottom of the bucket. You do not need to specify the size of the buoyant force, only how it changes with depth. ••

15. Glass A contains 250 g of water, and an identical glass B contains 220 g of water with a 30-g cube of ice floating in it. How do the water levels in the two glasses compare? ••

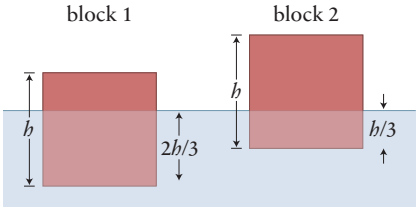
16. A rowboat floating in a small swimming pool has a large concrete block in it. If the block is thrown overboard into the pool, does the water level in the pool rise, fall, or stay the same? ••

17. You have constructed a raft that is 2.00 m long, 2.00 m wide, and 0.100 m thick. If the mass of the raft is 40.0 kg and your mass is 55.0 kg, how many 65.0-kg friends can cross a river with you during one trip? ••

18. A spherical balloon filled with helium gas must lift a 1.50-kg scientific payload off the ground, and the mass of the empty balloon is 0.500 kg. (*a*) What is the minimum diameter of the balloon? (*b*) A helium shortage forces the scientists to consider using nitrogen gas in the balloon. What minimum balloon diameter is needed if nitrogen gas is used? The mass densities are 1.286 kg/m^3 for air, 0.179 kg/m^3 for helium gas, and 1.251 kg/m^3 for nitrogen gas. ••

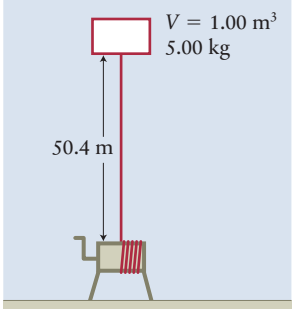
19. Two same-sized blocks of wood float in a tank as shown in Figure P18.19. Block 1 has two-thirds of its volume submerged, and block 2 has one-third of its volume submerged. Determine the ratio ρ_1/ρ_2 of the mass densities of the blocks. ••

Figure P18.19



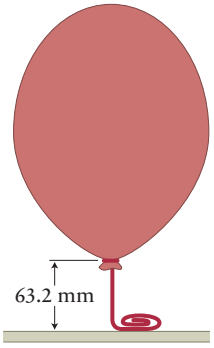
20. A 5.00-kg air-filled, sealed, rigid float tank that has a volume of 1.00 m^3 is pulled 50.4 m down to the seafloor in order to assist in lifting a sunken object. A diver standing on the seafloor cranks a winch to pull the tank down (Figure P18.20). (*a*) How much work is required to pull the tank down? (Ignore the mass of the air in the tank and the mass of the winch rope. Assume that the water mass density does not change with depth.) (*b*) What is the average power required if it takes 10 min to winch down the tank? ••

Figure P18.20



21. You want to determine the lifting force of a helium balloon. When you tie a long cord to the balloon and then release the balloon, it floats up until it supports a 63.2-mm length of cord and then it remains stationary, as shown in Figure P18.21. A 1.00-m length of the cord has a mass of 0.0160 kg, and the mass of the uninflated balloon is 2300 mg. What is the volume of the inflated balloon? The mass densities are 1.286 kg/m^3 for air and 0.179 kg/m^3 for helium. ••

Figure P18.21



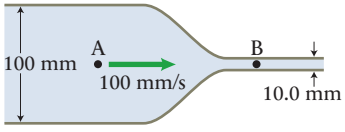
22. A cylindrical water tank that has a diameter of 2.50 m contains water to a depth of 3.75 m. A 40.0-kg object placed in the water floats with 75% of its volume submerged. By how much does the addition of the object increase the pressure in the water at the bottom of the tank? •••

23. An uninflated balloon of mass 6.00×10^{-3} kg has a volume of 6.90×10^{-2} m^3 when fully inflated with a mixture of helium and hydrogen gas. You tie a long, heavy cord to the balloon and allow the balloon to float above a table. The length of cord lifted off the table is 283 mm, and the linear mass density of the cord is 0.260 kg/m. (*a*) What is the mass density of the gas mixture in the balloon? (*b*) What is the ratio of helium to hydrogen in the mixture? The mass densities are 1.286 kg/m^3 for air, 0.179 kg/m^3 for helium gas, and 0.090 kg/m^3 for hydrogen gas. •••

18.3 Fluid flow

24. The pipe in Figure P18.24 shrinks from a diameter of 100 mm in the wide section to a diameter of 10.0 mm in the narrow section. If the water's speed is 100 mm/s at location A, what is its speed at location B? •

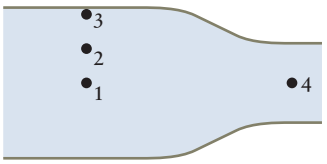
Figure P18.24



25. In an air duct running along the outside of a building, air flowing through the duct has a mass density of 1.20 kg/m^3 and is moving at 5.38 m/s. The duct passes through a wall into an air-conditioned room. If the mass density of the air in the duct rises to 1.24 kg/m^3 once the air reaches the air-conditioned region, what is the air's new speed? (Ignore the small dimensional change of the duct material in the colder room.) •

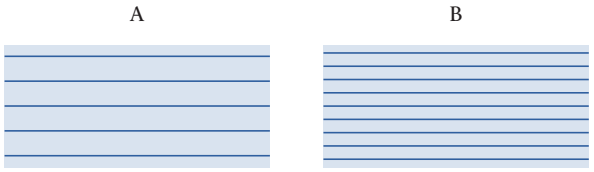
26. A viscous liquid is flowing through the pipe in Figure P18.26. If the flow is laminar, rank the speed of the liquid at points 1–4 in order of increasing speed. •

Figure P18.26



27. Figure P18.27 shows streamlines of water flowing in two regions A and B of the same pipe (neither of the two regions shown spans the entire diameter of the pipe). Between these two regions no other pipes are connected to this one. Compare the speed, pressure, and pipe diameter in the two regions A and B. ••

Figure P18.27



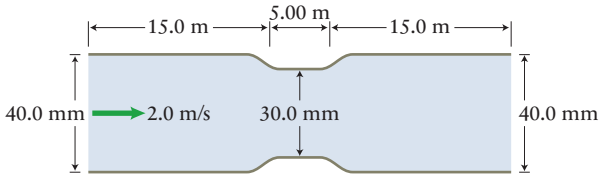
28. The water in a river flows gently in a continuous, straight channel of uniform width. Over a certain distance downstream, the water's speed decreases. What is most likely happening? State your reasons. ••

29. You are traveling in a car when the driver lights a cigarette. Why is it that when you open a window slightly the smoky air blows out of the car? Why doesn't air from the outside flow in through the opening instead? ••

30. The capped end of a pipe in your basement springs a leak. The pipe has a diameter of 50.0 mm and is attached to a joist so that it is 2.00 m above the basement floor. The water squirts out horizontally through a hole that has a diameter of 3.00 mm and hits the floor 6.00 m from the leak. What is the speed of the water inside the pipe as it moves toward the leak? Ignore air resistance. ••

31. A garden hose 30.0 m long and 40.0 mm in diameter attached to a faucet carries water that enters the hose at 2.00 m/s. The hose was previously cut at its center and has been repaired by splicing a length of hose that is 5.00 m long and 30.0 mm in diameter into the break (Figure P18.31). (*a*) What is the water speed in the small middle section of the repaired hose? (*b*) How long does the water take to flow the length of the repaired hose? ••

Figure P18.31



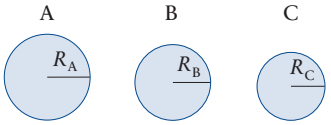
32. Consider a nonviscous fluid undergoing laminar flow in an open, horizontal channel. The Bernoulli effect says that where the flow speed is greater, the pressure is smaller, but the pressure at the surface, in contact with the surrounding air, is atmospheric pressure everywhere. How is that possible? •••

18.4 Surface effects

33. A drop of liquid is placed on a surface. A particle of the liquid at the contact point is subject to a cohesive force of 4.00 N directed 52.0° above the surface. If the contact angle is 104°, what is the magnitude of the adhesive force between the particle and the surface? •

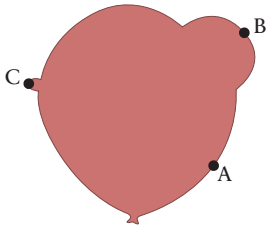
34. Three water droplets are shown in Figure P18.34. All three are in the same room with the same atmospheric pressure. In which of the three droplets is the pressure greatest and why? •

Figure P18.34



35. A partially inflated and irregularly shaped balloon is shown in Figure P18.35. Rank the points on the balloon's surface in order of (*a*) increasing tension in the membrane and (*b*) increasing pressure on the inside. •

Figure P18.35



36. Suppose a certain liquid in contact with a surface has cohesive and adhesive forces such that its behavior is exactly between wetting and not wetting a surface. How do the magnitudes of the adhesive and cohesive forces compare? ••

37. A glass capillary tube has a diameter of 2.00 mm and contains a small amount of liquid that forms a concave spherical meniscus. A particle of the liquid located where the liquid surface meets the tube wall is subject to a cohesive force of 4.00 N at an angle of 35.0° to the wall and an adhesive force of 1.00 N. What is the radius of curvature of the meniscus? ••

38. Dishwashing liquid dissolved in water forms an aqueous solution containing bubbles (the "suds") that are soap-film spheres filled with air. How does the pressure difference across the surface of one of these bubbles in the solution compare with the pressure difference across the surface of a bubble of the same size that escapes from the solution and floats in the air? ••

39. What is the relationship between the radius of a capillary tube R_{tube} and the radius of curvature R_{men} of the spherical meniscus a liquid forms in the tube, in terms of the contact angle θ_c? Answer for both a wetting liquid and a nonwetting liquid. (Hint: See Figure 18.36.) ••

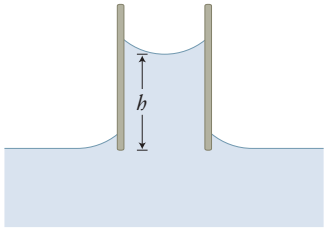
40. Liquid A forms a concave meniscus of radius R in a capillary tube in which its capillary rise is h_A. Liquid B has twice the surface tension and twice the mass density of liquid A and forms a concave meniscus of radius $R/2$ in an identical capillary tube. How does the capillary rise h_B of liquid B compare with h_A? ••

41. Figure 18.32 shows a drop of a wetting liquid and a drop of a nonwetting liquid sitting on a horizontal solid surface and surrounded by air. Show that, for both drops, the relationship between the contact angle and the magnitudes of the cohesive and adhesive forces exerted on a particle of liquid at the gas-liquid-solid junction is $2\sin(\theta_c/2) = F_{\text{cohesive}}/F_{\text{adhesive}}$. Assume that the surface is smooth and clean, the liquid contains no impurities, and the contact angle is not influenced by the gas. •••

42. You have three capillary tubes that are made of paraffin wax and have diameters of 1.00 mm, 2.00 mm, and 3.00 mm. (*a*) Is the meniscus that water forms in each tube convex or concave? (*b*) How might you change the water so that the answer to part *a* changes? (*c*) Because any liquid surface in a capillary tube curves either concavely or convexly, the peak or trough of the meniscus is vertically higher or lower than the location where the liquid surface meets the tube wall. Which of the three capillary tubes has the greatest vertical separation distance between the liquid-wall contact point and the meniscus peak or trough? •••

43. Figure P18.43 shows the capillary rise h for the liquid in a capillary tube sitting in a dish of liquid. Suppose you insert an identical capillary tube into the liquid but push this tube deep enough into the liquid to make the top of the tube be a distance $h/2$ above the surface of the liquid in the dish. How far do you expect the liquid to rise in this tube? •••

Figure P18.43



18.5 Pressure and gravity

44. A certain class of submarines can dive to 730 m below sea level. What is the pressure at this depth? •

45. The thresher shark has large eyes, about 0.10 m in diameter. Calculate the force exerted on a thresher shark's eye by the hydrostatic pressure in ocean water at a depth of 400 m. (Assume the water's mass density at this depth is 1000 kg/m^3.) •

46. Timbers used in building construction can withstand a pressure of 1.4×10^7 Pa before being crushed. If the mass density of wood is $\rho = 960$ kg/m^3, calculate the height h a pile of timbers would have to be in order for the ones at the bottom to be crushed. Assume that all the timbers have the same size and shape and that they are piled one on top of the other. ••

47. A cylindrical space capsule lands in the ocean. This capsule is 2.34 m long, 1.10 m in diameter, and weighted at one end so that it floats with its long central axis vertical and 0.820 m of its length above the water surface. The mass density of sea water is 1025 kg/m^3. (*a*) What is the magnitude of the buoyant force exerted on the capsule? (*b*) Suppose the capsule springs a leak and sinks. What is the magnitude of the buoyant force exerted on it once it sinks? (*c*) Why does the capsule sink if the magnitude of the buoyant force is greater after it has sunk than when it was floating? ••

48. A certain diving bell for use in the ocean has a square window measuring 0.250 m on each side. The inside of the bell is pressurized and maintained at a pressure of 2.00 atm. How deep can this bell be lowered before the vector sum of forces exerted by the seawater on the window exceeds 195 kN? The mass density of seawater is 1025 kg/m^3. ••

49. Pascal demonstrated his principle by causing a wooden barrel full of water to burst when a small amount of additional water was poured into a long, thin tube extending vertically upward from the top of the barrel. Explain why this worked and estimate the magnitude of the force that burst the barrel's top (diameter 0.42 m) and the length of the tube (diameter 6.0×10^{-3} m) if 3.39×10^{-4} m^3 of water (less than 12 fluid ounces) was added. ••

50. When a garden hose with an output diameter of 20 mm is directed straight upward, the stream of water rises to a height of 0.15 m. You then use your thumb to partially cover the output opening so that its diameter is reduced to 10 mm. How high does the water rise now? Ignore drag and assume that the smaller opening you create with your thumb is circular. ••

51. (*a*) For the garden hose of Problem 50, calculate how high the water would rise if you constricted the opening to make its diameter 1.0 mm rather than 10 mm. (*b*) Based on your experience, would you expect to see such a height from the stream? If not, what physical effects do you think are responsible? ••

52. (*a*) Consider a fluid whose mass density ρ varies with height y above sea level. Use Eq. 18.6, $P_1 + \rho g y_1 = P_2 + \rho g y_2$, to obtain an expression that shows how dP, the change in pressure along a column of the fluid of infinitesimal height dy, varies with y. (*b*) Using the simplification that the mass density of air at any given height y above sea level is proportional to atmospheric pressure at that height, integrate the differential equation you found in part *a* to obtain an expression that shows how atmospheric pressure varies with height y above sea level. •••

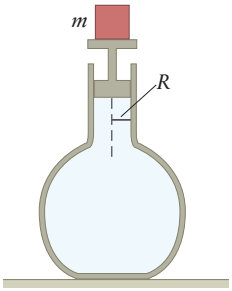
53. A river valley is in the shape of a V. A dam is built in this valley, so it also has a V shape, of height h and width w at the top and essentially zero at the bottom. Write an expression for the magnitude of the force that the water exerts against this dam. •••

18.6 Working with pressure

54. If the outside air pressure changes from 1.0 atm to 1.1 atm, by how much does mercury in a barometer rise? The mass density of mercury is 13,534 kg/m^3. •

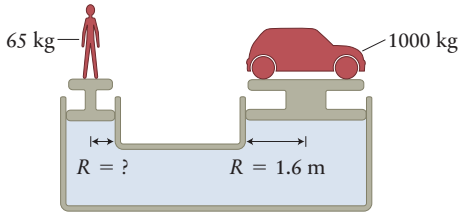
55. A small block of mass m sits atop a piston that seals a volume of air in a flask in which the radius of the neck is R (Figure P18.55). Obtain an equation for the pressure in the bottle as a function of atmospheric pressure. •

Figure P18.55



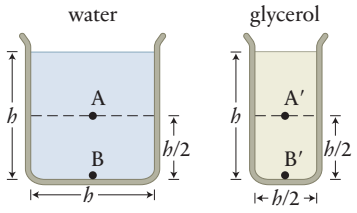
56. The two pistons in the hydraulic system shown in Figure P18.56 are coupled through a reservoir of water. The pistons are at the same height above ground. The piston on the right has a radius of 1.6 m, and the mass of the car sitting on the piston is 1000 kg. You are standing on the left piston, and neither you nor the car is moving. If your mass is 65 kg, what is the radius of the left piston? •

Figure P18.56



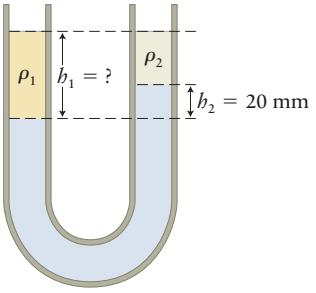
57. The container on the left in Figure P18.57 is a cube of side length h. The dimensions of the container on the right are height h, width $h/2$, and length $h/2$. Points A and A′ are at height $h/2$, and points B and B′ are at the bottoms of the containers. Calculate the ratio of gauge pressures $(P_A/P_B)/(P_{A'}/P_{B'})$ if the left container is filled with water and the right container is filled with glycerol. The mass density of glycerol is 1261 kg/m^3. ••

Figure P18.57



58. After pouring some water into the U-tube of Figure P18.58, you pour a volume V_1 of oil of mass density $\rho_1 = 800$ kg/m^3 into the left side of the tube and a volume V_2 of oil of mass density $\rho_2 = 700$ kg/m^3 into the right side. Once the two oils have been added, the levels on the left and right sides are the same, and the water level on the right side is a height $h_2 = 20$ mm higher than the water level on the left side. What is the height h_1 of the oil on the left side? ••

Figure P18.58



59. Initially, both legs of a mercury manometer are open to the atmosphere, and the mercury height in the left leg is 40.0 mm. You inflate a balloon until it has a surface area of 0.300 m^2 and affix its open end to the opening of the manometer's right leg. When you do this, the mercury level in the left leg changes to 65 mm. What is the magnitude of the inward force exerted by the balloon's elasticity on the air in the balloon? The mass density of mercury is 13.5×10^3 kg/m^3. ••

60. Suppose you increase the gauge pressure in the tires of your car by 10%. (*a*) How does the area of the tires in contact with a level road change? (*b*) If the gauge pressure in the tires is 2.0 atm before you increase it, by what percentage does the actual tire pressure change? ••

61. In a service station's hydraulic lift, the diameter of the large piston is 200 mm and the diameter of the small piston is 30 mm. Force is exerted on the small piston by an air compressor. (*a*) How much gauge pressure must the compressor supply to lift a 1500-kg car? (*b*) If the car must be raised 1.8 m off the floor, through what distance must the small piston travel? ••

62. A hydrogen-filled balloon is used to lift a 125-kg stone off the ground. The basket holding the stone has a mass of 15.0 kg. What must the minimum radius R of the balloon be in order to lift the stone off the ground? Use 1.29 kg/m^3 for the mass density of air, and ignore the mass of the hydrogen and the mass of the balloon. ••

63. In a large vat that is initially empty, a solid aluminum cylinder hangs by an ideal spring. The cylinder is 52.5 mm long and 30.0 mm in diameter, and the spring is stretched 42.0 mm. The vat is then filled with a liquid, submerging the cylinder and causing it to rise 11.8 mm. What is the mass density of the liquid? The mass density of aluminum is 2700 kg/m^3. ••

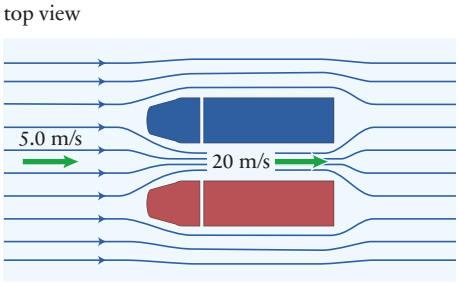
64. A 15.0-kg block measures 0.750 m top to bottom, and each horizontal face has an area of 0.0125 m^2. The block hangs, with its long central axis vertical, from a wire that can support a maximum tension of 135 N. The block is initially submerged in water, but then the water level is gradually lowered. What is the greatest distance the surface of the water can be below the top of the block without breaking the wire? ••

65. At an atmospheric pressure of 1.00 atm, the height of the liquid in the tube of a mercury barometer is 760 mm *only if* you ignore capillary rise in the tube. Obtain an expression that shows how h, the number of millimeters the mercury height differs from 760 mm, depends on the diameter d of the tube, and then calculate the value of h for $d = 7.0$ mm. The contact angle for mercury on this glass tube is 137°, and its mass density is 13.5×10^3 kg/m^3. •••
66. A small, smooth cube made of wood that has a mass density of 850 kg/m^3 is dropped from rest 5.00 m above the surface of a lake. Determine (*a*) the maximum depth the cube reaches and (*b*) the time interval the wood is under water before it returns to the surface. Assume that viscous drag is negligible so that the only force the water exerts on the cube is the buoyant force. •••
67. A right circular cone that has altitude h and a base of radius R is submerged in a liquid of mass density ρ. The cone's vertex points directly downward, so that the circular top face of the cone is horizontal. The volume of a right circular cone is $\pi R^2 h/3$. (*a*) By direct integration of the pressure over the surface of the cone, calculate the magnitude of the upward force exerted by the liquid on the cone. (*b*) Show that your answer in part *a* is consistent with Archimedes' principle. •••

18.7 Bernoulli's equation

68. Water from a faucet can fill a 4.00-L bucket in 1 min. If the diameter of the water pipe feeding the faucet is 12.5 mm, what is the speed of the water in the pipe? •
69. The water behind a dam built across one end of a large lake is 25.2 m deep. If the dam springs a leak 1.40 m above the lakebed, at what speed does water exit the hole? •
70. Two tractor trailers, each 16 m long and 4.0 m tall, are parked next to each other as shown in Figure P18.70. A light wind that is blowing at 5.0 m/s in the parking lot but at 20 m/s between the trailers causes the air pressure between them to drop. What is the magnitude of the force exerted by the air on each trailer? (Use 1.0 kg/m^3 for the air mass density.) •

Figure P18.70



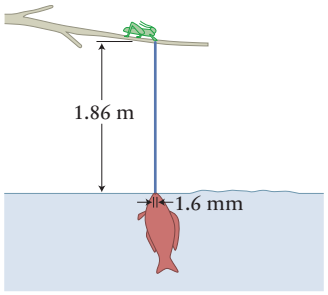
71. Water leaks out of a small hole in the side of a bucket. The hole is a distance d below the surface of the water, and the diameter of the hole is much smaller than the diameter of the bucket. Example 18.9 derived an expression for the speed v_2 at which water exits the hole by assuming that $v_1 \approx 0$ for the speed at which the water surface moves downward. Derive an expression for v_2 when v_1 cannot be ignored. ••
72. You carry a fire hose up a ladder to a height of 10.0 m above ground level and aim the nozzle at a burning roof that is 9.00 m high. You hold the hose horizontally and notice that the water strikes the roof at a horizontal distance of 7.0 m from where it exits the nozzle. The hose is connected to a large pressurized chamber in the fire truck 0.500 m above ground level. What is the pressure in the chamber? Ignore air resistance. ••
73. In an aorta that has a radius of 11 mm and a transmural pressure (pressure difference across the aorta wall) of 12 kPa, blood is moving in laminar flow at 350 mm/s. What percentage of the cross-sectional area of this aorta must be blocked in order for the transmural pressure in the blocked part to fall to zero? Assume that the blood has a mass density of 1060 kg/m^3 and is nonviscous, incompressible, and undergoing laminar flow at a speed of 350 mm/s in the unobstructed part. Ignore any pressure variation with height. ••
74. Does the laminar flow of a viscous fluid exhibit the Bernoulli effect? Does Bernoulli's equation (Eq. 18.36) apply to such a liquid? ••
75. A cylindrical tube that is 2.20 m long and has a radius of 150 mm is filled with water. It is oriented with its long central axis vertical, and it is open to the air at the upper end. A hole 80.0 mm in radius is drilled in the bottom, and the water is allowed to drain out. When the water level is half the height of the tube, (*a*) what is the speed of the water exiting though the hole, (*b*) at what rate (in meters cubed per second) is the water leaving the tube, and (*c*) how fast is the top of the water column falling? ••
76. Use integration to obtain an expression for the time interval needed for the water surface in the bucket in Problem 71 (or Example 18.9) to drain down to the level of the hole. You can't ignore v_1 anymore, but you can use either the expression for v_2 obtained in Problem 71 or an approximate expression for v_2 that is valid during most of the time interval in which the water is draining. •••

18.8 Viscosity and surface tension

77. A hollow stirrer that has a diameter of 1.0 mm is inserted into a cup of water at 20 °C. The surface tension of water at that temperature is $\gamma = 7.28 \times 10^{-2}$ N/m. What is the mass of the water supported by the surface tension in the stirrer? •
78. You insert a pipette of diameter 0.700 mm into a liquid that is at 20.0 °C and has a mass density of 1261 kg/m^3. If the liquid rises 29.0 mm up the pipette due to surface tension, what is the surface tension of the liquid? •
79. The volume flow rate of water at 20.0 °C through a cylindrical pipe is 2.00×10^{-4} m^3/s. What is the volume flow rate of castor oil moving through this pipe under the same conditions? At 20.0 °C, $\eta_{\text{castor oil}} = 0.986$ Pa · s. •
80. In a low-flow showerhead that contains 18 nozzle holes, the volume flow rate of the water is 4.00 L/min and the water speed is 4.30 m/s. What is the diameter of each hole? ••

81. In the device for measuring surface tension shown in Figure 18.58, the length of the movable wire is ℓ. When the film holding the wire in place is water at 20 °C, you must do 1.46 μJ of work on the wire to move it a distance Δx downward. If the film were instead mercury at the same temperature, how much work would be required to move the wire a distance $\Delta x/2$? The surface tensions at 20 °C are 7.28×10^{-2} N/m for water and 0.465 N/m for mercury. ••
82. A horizontal pipeline is 500 mm in diameter and carries oil at a rate of 150 kg/s. This oil has a mass density of 850 kg/m^3 and a viscosity of 0.224 Pa·s at the temperature of the pipeline. If the pressure in the oil is 3.75 atm at the front end, how long can the pipeline be so that the pressure at the far end is no less than 2.50 atm? ••
83. Water from a faucet fills a bathtub with 100 L of water in 5.00 min. That water travels through a cylindrical pipe that has a diameter of 12.5 mm. What is the maximum speed of the water? ••
84. The archerfish can form a tube by curling its tongue against a groove in the top of its mouth and then snapping shut its gills. The fish can then spit a stream of water at its prey (Figure P18.84). One archerfish was observed shooting a stream of water to a height of 1.86 m, and his tongue tube was measured to be 1.60 mm in diameter. What is the minimum volume flow rate of the spit water? ••

Figure P18.84



85. Two horizontal cylindrical pipes are connected together, and a viscous liquid undergoes laminar flow through them. The first pipe has radius R_1 and length ℓ_1, and the second has radius R_2 and length $2\ell_1$. If the pressure where the pipes connect is midway between the pressures at the other ends of the pipes, how does R_2 compare with R_1? ••
86. A liquid of viscosity 0.456 Pa·s and mass density 882 kg/m^3 flows through a pipe at a volume flow rate of 2.00 m^3/s. You need to increase this rate to 8.50 m^3/s by adding to the liquid a substance that will change its viscosity but not appreciably affect its mass density. (*a*) If nothing else (including the pressure difference) changes, what new viscosity value do you need in order to achieve the desired volume flow rate? What is the *mass* flow rate, in kilograms per second, (*b*) before you change the viscosity and (*c*) after you make the change? ••
87. Drops of an oil that has a mass density of 700 kg/m^3 are released from a broken undersea oil pipe that is 1000 m below the ocean surface. (*a*) If the drops have an average diameter of 100 μm, how long do they take to rise to the surface? Hint: The drops quickly reach terminal speed and experience a viscous drag force of magnitude $F^{\text{d}}_{\text{wo}} = (0.0200 \text{ Pa} \cdot \text{s})Rv$, where R is the radius of the drops and v is their speed. (*b*) After a chemical dispersant is applied to the oil to reduce the average drop diameter to 2.0 μm, how long do the drops take to rise to the surface? Use 1000 kg/m^3 for the mass density of ocean water. •••
88. Consider the cylindrical outer surface of the column of liquid in a tube of length ℓ and radius R. Imagine slicing this surface into two half cylinders along the long central axis of the pipe. Show, by integration, that Eq. 18.52 is valid—that is, (*a*) that the magnitude of the force due to the pressure difference $P_{\text{in}} - P_{\text{out}}$ across the right semi-cylindrical surface (balanced by the force due to the surface tension due to the left half) is equal to the pressure difference $(P_{\text{in}} - P_{\text{out}})$ times $2\ell R$, and (*b*) that the direction of this force is perpendicular to the pipe length and bisects the semi-cylindrical surface. •••

Additional Problems

89. A liquid moves through a pipe of diameter d at volume flow rate Q_1. When the pipe diameter is reduced to $d/4$ while the pressure is held constant, what is the new volume flow rate Q_2 in terms of Q_1? •
90. A blood vessel has a diameter of 3.20 mm because it is partially clogged by cholesterol. You want to remove enough cholesterol to double the volume flow rate of the blood through the vessel without changing its speed. What should be the new vessel diameter? Assume the blood is nonviscous and the flow is laminar. •
91. Europa, a satellite of Jupiter, appears to have a liquid ocean of water below an ice layer on its surface. The acceleration due to gravity at the surface is 1.3 m/s^2, and there is no atmosphere. You are designing a probe to descend into Europa's ocean, and it must withstand the pressure 125 m deep in that ocean. When you are testing the probe here on Earth, how deep in Earth's ocean should you lower the probe to have it experience the same pressure it will experience when it is 125 m down in Europa's ocean? The mass density of seawater is 1025 kg/m^3, and assume the water in Europa's ocean has a mass density of 1000 kg/m^3. •
92. Estimate how much the air surrounding your body affects the scale reading when you stand on a spring scale. ••
93. In an automobile braking system, the brake pedal pushes on one end of a column of liquid (the *brake fluid*) in a rigid tube. At the other end of the tube, the liquid presses on a piston attached to a brake shoe that pushes against the brake drum and slows the car's wheels by friction. In one such system, the diameter of the tube at which the brake pedal pushes is 12.0 mm, the diameter of the piston at the brake drum is 44.0 mm, and the mass density of the liquid is 850 kg/m^3. The coefficient of static friction between the brake shoe and the brake drum is 0.880, the coefficient of kinetic friction between them is 0.650, and the piston and brake shoe push perpendicular to the brake drum. When you push on the brake pedal with a force of 125 N, what are (*a*) the magnitude of the normal contact force exerted by the brake shoe on the drum and (*b*) the magnitude of the frictional contact force exerted by the brake shoe on the drum? ••

94. A thick layer of oil of mass density 800 kg/m^3 floats atop (but does not mix with) water in a deep container. You gently place a sphere of uniform mass density in the container so that it floats with 65.0% of its volume in the water and the remaining 35.0% in the oil. What is the mass density of the material of which the sphere is made? ●●

95. In bubble tea with pearls, a popular drink from Taiwan, balls of tapioca are put into cold tea mixed with sugar and milk. Some of the balls fall to the bottom of the cup, others rise to the surface, and a few hang motionless in the liquid. (*a*) If the tea mixture has a mass density of 1350 kg/m^3, what is the approximate mass of a ball that neither rises nor falls and just fits into a straw that has a diameter of 10 mm? (*b*) Do you expect the ball to continue to hang motionless if the temperature of the mixture changes? ●●

96. A cubical box measures 1.50 m on each side and has a mass of 3000 kg. Its walls are made of a dense but very thin metal, and the box is placed upright in a lake. (*a*) Once equilibrium is achieved, how far below the lake surface is the bottom of the box? (*b*) If you slowly add water to the box, how deep can the water level in the box be before the box sinks? Assume that the box always remains level. ●●

97. On a planet the same size as Earth, ocean water has a mass density of 1000 kg/m^3. A space probe measures the atmospheric pressure at the surface of this planet to be 2.40×10^5 Pa. At a depth of 150 m in the ocean, the probe reports a pressure of 14.4×10^5 Pa. What is the mass of the planet? ●●

98. A fire hose shoots water from ground level at an angle of 60.0° above the horizontal, with negligible air resistance. The initial speed of the water is such that the water strikes a building horizontally at a height of 25.0 m above the ground. The diameter of the cylindrical hose is 0.120 m at the location where the water exits. What are (*a*) the volume flow rate of water leaving the hose and (*b*) the mass flow rate of that water in kilograms per second? ●●

99. A spherical measuring device of radius 250 mm and mass density 1500 kg/m^3 is suspended under water by two wires. Wire A pulls horizontally to the left on the sphere, and wire B pulls upward and to the right at an angle of 50.0° above the horizontal. Calculate the tension in each wire. ●●

100. A long *evacuated* tube has a cross-sectional area of 1.50×10^{-3} m^2 and is permanently sealed at one end but can be opened at the other. When the tube is opened while oriented vertically with its open end just below the surface of the liquid in a large open beaker, the liquid rises to a height of 1.28 m in the tube. If the air pressure in the room is 1.00 atm, what is the mass density of the liquid? ●●

101. A large iceberg in the shape of a cube has a mass density of 917 kg/m^3 and floats upright in seawater that has a mass density of 1024 kg/m^3. (*a*) If the sides of the cube measure 75.0 m, what is the iceberg's frequency of oscillation? Hint: First show that the motion is simple harmonic. (*b*) What would be the frequency of a cube-shaped iceberg that was twice as massive? Solve this part by proportional reasoning. ●●

102. A cylindrical log of length ℓ and radius R floats horizontally half submerged in a liquid of mass density ρ. (*a*) By direct integration of the pressure over the surface of the log, calculate the vector sum of the upward forces exerted on the log by the liquid. (*b*) Show that your answer in part *a* is consistent with Archimedes' principle. ●●●

103. Working on a team designing a new cargo ship, you have been charged with investigating what dimensions enable a steel ship to carry as much cargo as possible. Certainly the larger the ship, the more cargo it can carry, but you are given a fixed amount of steel with which to work. Assume the ship can be modeled as a box of any dimensions. You realize at once that you cannot use the design with the smallest mass density. ●●● CR

104. Watching a parent and child balancing on a seesaw, you begin to wonder what effect the buoyant force exerted by the surrounding air might have on the distance each of them must be from the pivot to achieve balance. ●●● CR

105. You have been given the task of determining, for a redwood tree 100 m tall, the mechanisms that allow the tree to grow to such a height and the mechanisms that prevent it from growing any taller. Using your knowledge of fluid dynamics, you determine that capillary rise causes water to rise a height h through a tree's xylem (the hollow tubelike cells in trees that carry water), and for a xylem radius r, this height is approximately given by $h = (0.137\ \text{J/m}^2)/(\rho_{\text{water}} g r)$. Using a microscope, you observe that the xylem has a diameter of about 5 μm. You are convinced that a mechanism other than capillary rise is needed to enable the water to rise to the top leaves. ●●● CR



19

Entropy

CONCEPTS

19.1 States
19.2 Equipartition of energy
19.3 Equipartition of space
19.4 Evolution toward the most probable macrostate

QUANTITATIVE TOOLS

19.5 Dependence of entropy on volume
19.6 Dependence of entropy on energy
19.7 Properties of a monatomic ideal gas
19.8 Entropy of a monatomic ideal gas

The conservation laws allow us to analyze a broad range of phenomena, yet they cannot explain the *irreversibility* of time. Cars are damaged in a collision, a cube of ice melts when placed in a drink, and a drop of ink diffuses throughout a glass of water. These phenomena are *irreversible*, meaning they happen in one direction only. Always, no exceptions.

Yet there's nothing in the conservation laws that prohibits cars damaged in a collision to spontaneously move apart "undamaged," that prohibits an ice cube from forming in a glass of water, or that prohibits a collection of diffused ink molecules from coming together into a single drop. So why do phenomena such as these never happen? What is it that causes the irreversible flow of time toward the future?

As we shall see in this chapter, the underlying reason has to do with the randomness of the interactions at the atomic level. To help us quantify the tendency of systems to evolve in one direction only over time, I introduce a quantity, *entropy*, whose change is strongly connected to our sense of irreversibility.

19.1 States

Have you ever seen a motionless pendulum suddenly start swinging of its own accord? No; when no forces are applied, pendulums slow down and eventually come to rest. To understand why a swinging pendulum always slows down and never speeds up, let us consider a pendulum swinging back and forth inside a container. As the pendulum swings, kinetic energy gets converted to gravitational potential energy, which then gets converted to kinetic energy, which then . . . . If no energy is dissipated, the sum of these two energies remains constant and the pendulum keeps swinging forever.

Over time, however, the swings get smaller because with each swing a small part of the coherent mechanical energy of the pendulum is converted to incoherent thermal energy (see Section 7.3). This conversion is due to friction in the pendulum suspension and to collisions with the air molecules surrounding the pendulum. When the pendulum has come to rest, all of its mechanical energy has been dissipated. This energy is not gone—it still exists but now in the form of incoherent thermal energy associated with the motion of the molecules in the air surrounding the pendulum. So the universal tendency for the mechanical energy to dissipate appears to be related to the incoherent nature of the motion of atoms and molecules at the atomic scale.

Not one of the conservation laws we have studied forbids the conversion of that thermal energy back to mechanical energy. The pendulum could start swinging again if the air molecules would supply the kinetic energy. To understand why this transfer of energy from air molecules to pendulum is never observed, we need to look at molecular motion on the atomic scale. Let's begin by getting some feel for how much energy the pendulum and the air molecules have.

EXERCISE 19.1 Pendulum energy

Consider a 0.10-kg pendulum swinging at a maximum speed of 0.80 m/s inside a box that contains 1.0×10^{23} nitrogen molecules. The mass of a nitrogen molecule is 4.7×10^{-26} kg, and at room temperature a typical nitrogen molecule moves at 500 m/s. What are (*a*) the mechanical energy of the pendulum, (*b*) the average kinetic energy of one nitrogen molecule, and (*c*) the sum of the average kinetic energies of all the nitrogen molecules?

SOLUTION (*a*) The mechanical energy of the pendulum is equal to its maximum kinetic energy: $\frac{1}{2}mv^2 = \frac{1}{2}(0.10\text{ kg}) \times (0.80\text{ m/s})^2 = 3.2 \times 10^{-2}$ J. ✔

(*b*) The average kinetic energy of a nitrogen molecule is $\frac{1}{2}(4.7 \times 10^{-26}\text{ kg})(500\text{ m/s})^2 = 5.9 \times 10^{-21}$ J. Comparing this value with the result I obtained in part *a*, I see that the mechanical energy of the pendulum is more than 18 orders of magnitude (a billion billion times) greater than the average kinetic energy of the individual nitrogen molecules. ✔

(*c*) The sum of the average kinetic energies of all the nitrogen molecules is $(1.0 \times 10^{23})(5.9 \times 10^{-21}\text{ J}) = 5.9 \times 10^{2}$ J. Note that this thermal energy (the kinetic energy associated with the incoherent motion of all the nitrogen molecules) is more than four orders of magnitude greater than the mechanical energy of the pendulum. ✔

19.1 The nitrogen molecules in Exercise 19.1 move considerably faster than the pendulum and continuously bombard the pendulum from all sides. (*a*) Why does this bombardment slow the pendulum down? (*b*) When the pendulum is at rest, why don't we see the effects of this continuous bombardment?

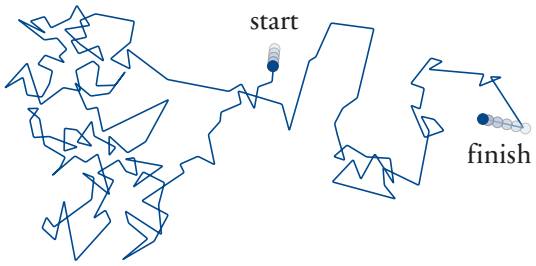
The inertia of a pendulum at rest is so much greater than that of the surrounding air molecules that the effect of collisions between air molecules and the pendulum is hardly noticeable. To see the effect of such collisions, we would have to greatly reduce the inertia of the pendulum. In 1827 the Scottish botanist Robert Brown was the first to observe the effect of this random molecular bombardment. Using a microscope to observe grains of pollen suspended in water, Brown noticed that the grains bounced around, following a zigzag path like the one shown in Figure 19.1 on the next page.

This random motion, called *Brownian motion*, is due to the random bombardment of the grains by the surrounding water molecules. Because the grains are small, the bombardment is not always equal on all sides, causing the grains to zigzag. The small random movements are what makes up Brownian motion.

19.2 How would Brownian motion change if the mass of the grain in Figure 19.1 were increased?

CONCEPTS

Figure 19.1 A grain of pollen suspended in water undergoes a random zigzag motion due to collisions with surrounding water molecules.



An important conclusion we can draw from the observation of Brownian motion in pollen grains suspended in water is that the bombardment is incoherent. As a result of this lack of coherence, the displacement of each grain over a long time interval is nearly zero and the grain never gains or loses any kinetic energy. If the grain were constantly kicked in the same direction, it would keep accelerating in that direction and gain kinetic energy. The same is true for a pendulum suspended in the air and hanging at rest: If it were constantly hit from one direction, it *would* start swinging.

Because it is impossible to treat the details of the incoherent atomic motion quantitatively, we need to use a very different approach to connect the atomic world of atoms and molecules to the *macroscopic* world we have described so far (*macroscopic* means "large relative to the size of a molecule"). In particular, we need to use probability theory to determine the likelihood of certain events taking place. As we shall see, the odds strongly favor slowing down a pendulum over speeding it up. To begin, let's review the essentials of probability theory using the throwing of two dice as an example.

CONCEPTS

EXERCISE 19.2 Throwing dice

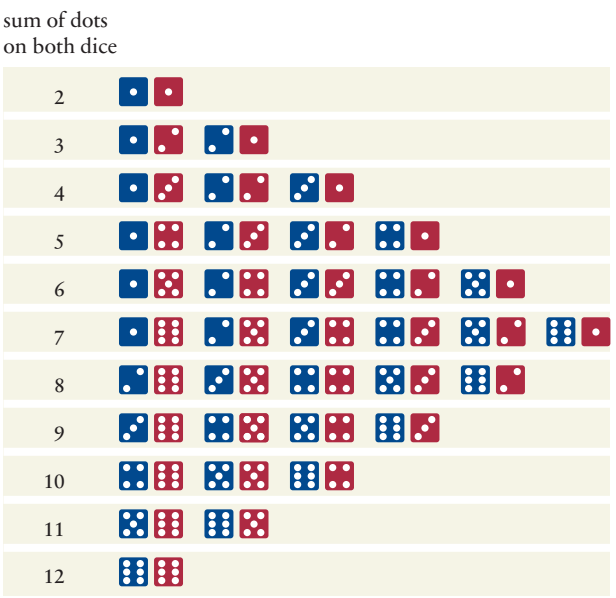
When you throw two dice, one red and one blue, many times, what is the fraction of throws for which the dots on the two dice add to 4?

SOLUTION There are 36 combinations for the pair of dice (**Figure 19.2**), three of which add to 4. On every throw, each of the 36 combinations is equally likely to occur. If you throw the dice many times, the fraction of throws that add to 4 approaches the ratio $3/36 = 1/12$. ✔

19.3 What sum of dots are you most likely to throw with two dice?

The **probability** that a certain event will occur is the fraction of times that event occurs in a large number of repetitions. Exercise 19.2 tells you that the probability of throwing a 4 is $1/12$. If you throw two dice 12 times, you won't necessarily throw a 4 exactly once. The more times you throw the dice, however, the closer the fraction of times you throw a 4 will be to $1/12$.

Figure 19.2 The 36 possible combinations for two dice.



If we consider the two dice as a system, there are several ways we can describe the state of this system. We can characterize it by the sum of the dots shown on the two dice after any given throw, or we can give a more detailed description by specifying the number of dots showing on each die. For systems with very large numbers of particles, we often care only about large-scale properties, not the detailed specifications of each particle. When we describe a system in terms of its large-scale properties, we are describing the system's **macrostate.** When we give a detailed specification of all the particles that make up the system, we are describing the system's **basic state.*** For a volume of gas, for example, the macrostate is described by large-scale properties that can be readily measured, such as temperature, pressure, and volume. The basic state is described by specifying the position and velocity of each individual gas particle.

Before describing a gas, let us return to the simpler system of two dice. The macrostate of the system after any throw is given by the sum thrown; the basic state is given by the number of dots shown on each die. As you can see in Figure 19.2, there are 36 basic states for the system and 11 macrostates. Because any basic state is equally likely to occur, the probability of any one basic state occurring is $1/36$. To calculate the probability of, say, the

*Often the term *microstate* is used instead of *basic state*. Because *microstate* and *macrostate* sound alike and are often hard to tell apart in spoken language, we'll use the term *basic state*.

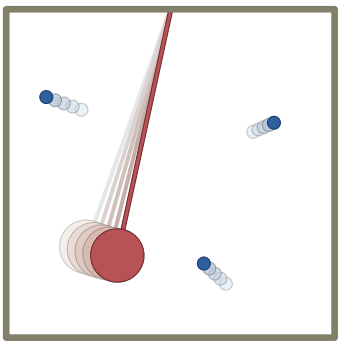
macrostate 9, we divide the number of basic states associated with this macrostate by the number of basic states associated with all the macrostates. Four out of the 36 basic states correspond to the macrostate 9. Therefore the probability of throwing a 9 is $(1/36)(4) = 4/36$. Because there are more basic states for the macrostate 7 than there are for the macrostate 9, you are more likely to throw, on average, a 7 than a 9.

19.4 Suppose you are tossing four coins and decide to define the system's macrostate after any throw as the number of heads thrown. How many (*a*) macrostates and (*b*) basic states are there for this system? (*c*) What is the probability of tossing three heads? (*d*) What is the most likely macrostate after any toss?

Let us now see how all this applies to a swinging pendulum. Consider one swinging inside a box containing three gas particles (**Figure 19.3**). The particles are *distinguishable*—that is, we can tell them apart—and they collide randomly, but elastically, with one another, with the walls of the box, and with the pendulum. According to *quantum mechanics*, the energy of confined systems can only change by very small, indivisible, discrete units called *quanta* (singular: *quantum*), and many experiments confirm this discrete nature of energy units. The collisions in the box thus transfer energy units between the colliding objects.

To get a feel for how energy units are redistributed as the particles collide with the pendulum, we ignore any dissipation of energy in the pendulum's suspension and we assume that there are six identical (and therefore

Figure 19.3 As a pendulum in a box swings, it collides with three particles. The pendulum slows down as its energy is redistributed into the particles.



indistinguishable) energy units to be distributed among the pendulum and the three particles. (We'll drop the restriction that the energy units are indistinguishable later.) We further assume that the collisions distribute the energy units randomly throughout the system: that is, the redistribution jumbles up both the speeds of the particles and the directions in which they move. This assumption is justified not only by the observation of Brownian motion but also by the fact that the assumption leads to a behavior that accurately describes our observations.

Let us investigate how the six energy units can be distributed among the pendulum and the three gas particles. If we describe the system's macrostate by the number of energy units in the pendulum, there are seven macrostates. **Figure 19.4** shows how, for each of these seven macrostates, the energy units can be distributed over the pendulum and the three particles. Note that as the

Figure 19.4 The ways in which six energy units can be distributed among the pendulum and three particles of Figure 19.3. Distribution of six energy units among pendulum and three particles; each solid-colored box corresponds to one of the six units of energy. The number of basic states associated with all macrostates is 84. As the number of energy units in the pendulum decreases (left), the fraction of basic states associated with that macrostate increases (right).

Number of energy units in pendulum	Number of energy units in particles	Number of basic states	Fraction of basic states
6	0 — 1	1	$\frac{1}{84}$
5	1 — 3	3	$\frac{3}{84}$
4	2 — 3 — 3	6	$\frac{6}{84}$
3	3 — 3 — 6 — 1	10	$\frac{10}{84}$
2	4 — 3 — 6 — 3 — 3	15	$\frac{15}{84}$
1	5 — 3 — 6 — 6 — 3 — 3	21	$\frac{21}{84}$
0	6 — 3 — 6 — 6 — 3 — 3 — 6 — 1	28	$\frac{28}{84}$

three basic states having one energy unit in one particle

amount of energy in the particles increases, the number of basic states associated with that macrostate increases.

19.5 What are the six ways in which you can distribute three energy units over three particles such that one of the particles has two energy units?

Because collisions completely randomize the energy distribution, each basic state is equally likely to occur, and we can use statistics to determine the probability of any given macrostate. The probability of any one basic state is 1/84 (because there are 84 equally likely basic states). The probability of any macrostate is the number of basic states associated with that macrostate divided by the number of basic states associated with all macrostates, which is the number given in the last column of Figure 19.4. **Figure 19.5** shows how this probability depends on the fraction of energy in the pendulum.

19.6 What is the probability of finding (*a*) all of the energy in the pendulum and (*b*) either zero or one unit of energy in the pendulum?

We "digitized" the energy by arbitrarily dividing it into six units. However, collisions could redistribute smaller units of energy, and so the mechanical energy of the pendulum could change by smaller amounts. **Figure 19.6** shows how the probability of finding a certain fraction of energy in the pendulum relative to that of finding zero energy in it changes as we divide the energy into smaller units. Note by comparing Figures 19.5 and 19.6 that the division into finer units has little effect on the relative probabilities. In other words, the picture we obtain by dividing the energy into very coarse energy units is a remarkably good approximation of the much more realistic fine division predicted by quantum mechanics.

Figure 19.5 As the fraction of energy in the pendulum increases, the probability of that corresponding macrostate decreases.

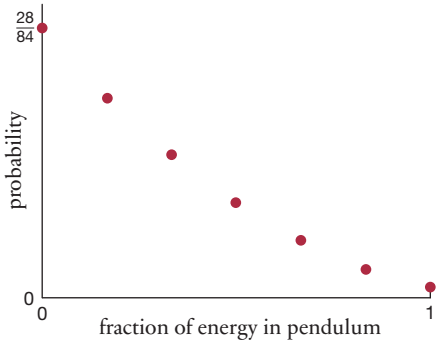
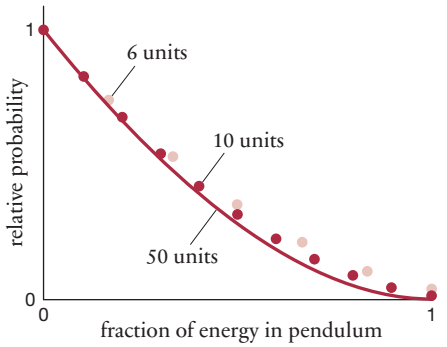


Figure 19.6 As the energy is divided into more units, the probability of finding a certain fraction of the system's energy in the pendulum relative to that of finding zero energy in it approaches the solid red line.



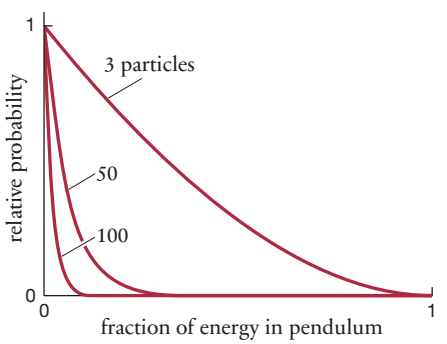
What happens if we increase the number of particles? **Figure 19.7** shows that as the number of particles increases, the spike in the relative probability near the origin sharpens and the relative probability of all 50 energy units in the pendulum drops precipitously: With 3 particles, it is on the order of 10^{-4}; with 50 particles, it is on the order of 10^{-29}; with 100 particles, it drops to 10^{-40}.

Even with a very large number of particles, however, the relative probability of the pendulum containing all the energy is nonzero. So, if we release a pendulum in a box that contains many particles and the mechanical energy of the pendulum is gradually transferred to the particles, there is a very small but nonzero probability of the pendulum gaining all its energy back again.

We can get a feeling for how long it would take for this to happen by multiplying the average time interval between collisions by the number of basic states for the system. This gives us the length of time it would take on average to cycle randomly through all the basic states. This time interval is called the *recurrence time* of the system.

In a box of volume 1 m^3 containing 3 air molecules at room temperature, the average time interval between

Figure 19.7 As the number of particles increases from 3 to 100, the probability of finding a certain fraction of the system's 50 energy units in the pendulum relative to that of finding zero energy in it decreases significantly.



CONCEPTS

collisions is approximately 10^{15} s. So the recurrence time is approximately $(84)(10^{15}\text{ s}) = 10^{17}$ s. If there are 10^{20} particles, the average time interval between collisions is approximately 10^{-19} s. For a large number of particles, the number of basic states is equal to the number of particles raised to the number of energy units: $(10^{20})^6 = 10^{120}$ (Section 19.6). Thus the recurrence time is approximately $(10^{120})(10^{-19}\text{ s}) = 10^{101}$ s. The age of the universe is only about 10^{18} s, so don't hold your breath waiting for the pendulum to recover all its energy!



19.7 (*a*) Averaged over a long time interval, for what fraction of the time interval does the pendulum in Figure 19.3 have more than 50% of the energy of the system? (*b*) Does this fraction increase, decrease, or stay the same as we increase the number of particles?

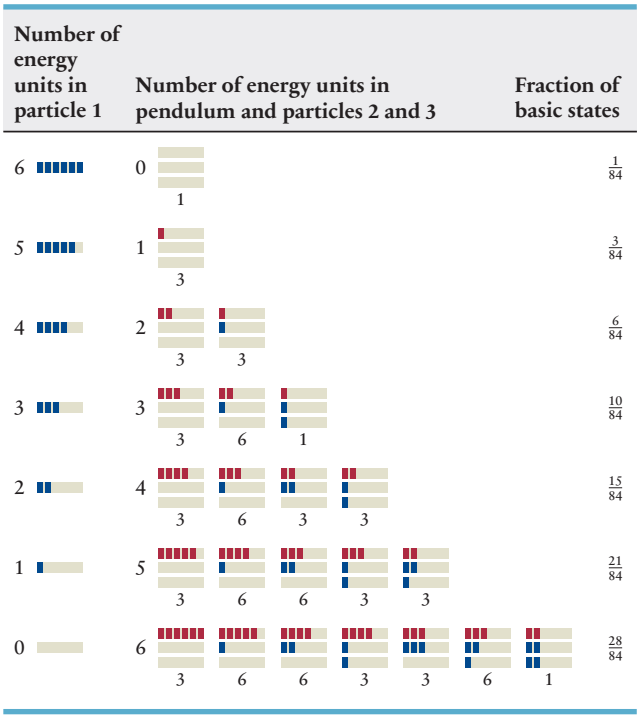
19.2 Equipartition of energy

In our description of a pendulum swinging among gas particles, we saw that as the number of particles increases, the probability of nonzero energy in the pendulum goes to zero (Figure 19.7). Suppose now that we have a gas of four particles instead of a pendulum and three particles. The bars for the energy units in the pendulum in Figure 19.4 then represent the energy of the fourth particle, and we see that the probability is highest for that particle to have little or no energy. We can use the same argument for each of the four particles: Each particle tends to have the smallest amount of energy possible. Because the thermal energy E_{th} associated with the incoherent motion of the particles is fixed, the system tends to distribute this energy equally over all the particles, with each particle having, on average, an amount of energy E_{th}/N, where N is the number of particles. If at any instant one of the particles has a greater fraction of the thermal energy E_{th}, this energy is rapidly redistributed over the rest of the system.

Each particle is like the pendulum. Therefore if we exchange the role of the pendulum with the role of one of the particles in Figure 19.4, we obtain the macrostates and basic states shown in **Figure 19.8**. Again, the odds are highest for that particle to have little or no energy, and as we saw in Figure 19.7, these odds become even more pronounced for large numbers of particles. Each particle tends to minimize its energy.

The situation is the same for the pendulum and the three particles. On average, each part of this system has an equal share $E/4$ of the system's energy E. If we replace either the pendulum or one of the particles with an oscillating spring, with a heavier particle, or with some other object, we can go through the same arguments and in each case we see that:

Figure 19.8 If we interchange the pendulum with one of the particles in Figure 19.4, we obtain an identical configuration of macrostates and basic states.



As long as the interactions between different parts of a system randomize the distribution of energy, each part of the system tends to have an equal share of the system's energy.

This equal distribution of energy among the parts of a system is called **equipartition of energy**. In Section 20.1 we shall define more precisely what "part of the system" means. For now, consider each indivisible element of the system—a particle, a pendulum, an oscillating spring, whatever—as a "container" for energy.

CONCEPTS

EXAMPLE 19.3 Pendulum dissipation

Consider again the 0.10-kg pendulum of Exercise 19.1. The pendulum starts out swinging at a maximum speed of 0.80 m/s and stops swinging in 5.0 min. It is in a box that contains 1.0×10^{23} nitrogen molecules, each having a mass of 4.7×10^{-26} kg and moving at 500 m/s. The pendulum collides, on average, 1.0×10^{22} times per second with the nitrogen molecules. (*a*) In one collision what amount of mechanical energy is transferred on average to a nitrogen molecule, and what is the ratio of this transferred energy to the average kinetic energy of the nitrogen molecule? (*b*) By what fraction has the average kinetic energy of a nitrogen molecule increased when the pendulum stops swinging? (*c*) If the final mechanical energy of the pendulum is equal to the average kinetic energy of a nitrogen molecule, what is the pendulum's average final speed?

(Continued)

1 GETTING STARTED The collisions gradually transfer all the pendulum's initial kinetic energy to the nitrogen molecules. For simplicity, I assume that each collision transfers the same amount of energy.

2 DEVISE PLAN The number of collisions that occur during the 5.0 min it takes the pendulum to stop swinging is $(3.0 \times 10^2\ \text{s})(1.0 \times 10^{22}\ \text{collisions/s}) = 3.0 \times 10^{24}$ collisions. Dividing the initial kinetic energy of the pendulum (3.2×10^{-2} J; Exercise 19.1*a*) by this number gives me the average energy transferred in each collision. I can then compare this energy with the average kinetic energy of a nitrogen molecule calculated in Exercise 19.1*b* (5.9×10^{-21} J).

When the pendulum stops, all its mechanical energy has been transferred to the molecules. In Exercise 19.1, I determined the initial mechanical energy of the pendulum (3.2×10^{-2} J) and the sum of the average kinetic energies of all the nitrogen molecules (5.9×10^2 J). The ratio of these quantities gives me the fractional increase in the kinetic energy of each molecule.

The final average kinetic energy of the pendulum is half of its final mechanical energy: $\frac{1}{2}m_{\text{pend}}(v^2_{\text{pend,f}})_{\text{av}} = \frac{1}{2}E_{\text{pend,f}}$. To calculate the pendulum's average final speed, I set $E_{\text{pend,f}}$ equal to the final average kinetic energy of a nitrogen molecule from Exercise 19.1*b*.

3 EXECUTE PLAN (*a*) The average amount of mechanical energy transferred from pendulum to molecule in each collision is $(3.2 \times 10^{-2}\ \text{J})/(3.0 \times 10^{24}\ \text{collisions}) = 1.1 \times 10^{-26}$ J/collision. From Exercise 19.1*b*, the average kinetic energy of each molecule is 5.9×10^{-21} J. So the ratio of the energy transferred per collision to the average molecular kinetic energy is $(1.1 \times 10^{-26}\ \text{J})/(5.9 \times 10^{-21}\ \text{J}) = 1.9 \times 10^{-6}$. ✔

(*b*) The fractional increase in the average kinetic energy of each nitrogen molecule is the initial kinetic energy of the pendulum (which all ends up in the molecules) divided by the combined kinetic energy of the molecules from Exercise 19.1: $(3.2 \times 10^{-2}\ \text{J})/(5.9 \times 10^2\ \text{J}) = 5.4 \times 10^{-5}$. ✔

(*c*) Because the final mechanical energy of the pendulum is equal to the average kinetic energy of a nitrogen molecule, I have $\frac{1}{2}m_{\text{pend}}(v^2_{\text{pend,f}})_{\text{av}} = \frac{1}{2}(5.9 \times 10^{-21}\ \text{J})$, or $(0.10\ \text{kg})(v^2_{\text{pend,f}})_{\text{av}} = 5.9 \times 10^{21}$ J. Therefore the final average speed of the pendulum is $(v_{\text{pend,f}})_{\text{av}} = [(5.9 \times 10^{-21}\ \text{J})/(0.10\ \text{kg})]^{1/2} = 2.4 \times 10^{-10}$ m/s. ✔

4 EVALUATE RESULT Part *a* tells me that each collision transfers only a very small fraction of the pendulum's energy. This makes sense because it takes many collisions for the pendulum to come to rest. The small value I obtained in part *b* for the fractional increase in the molecules' energy makes sense because I know from experience that the air temperature (a quantity related to the kinetic energy of the molecules in the air) in the region around a swinging pendulum does not increase appreciably as the pendulum moves. Finally, my answer to part *c* shows that, for all practical purposes, the final average speed of the pendulum indeed is zero, as I had assumed in part *a*.

As Example 19.3 illustrates, the pendulum begins with much more mechanical energy than each individual gas molecule. Each collision between the pendulum and a nitrogen molecule can increase or decrease the energy of the pendulum. On average, however, these energy exchanges gradually decrease the energy of the pendulum until all nitrogen molecules and the pendulum have roughly the same energy. The pendulum's excess energy has been dissipated over all the molecules, raising their average energy by a minuscule amount.

19.8 Suppose a container holds a gas made up of two kinds of particles 1 and 2, with $m_2 = 2m_1$. Assume the particles interact via elastic collisions. (*a*) How do the average kinetic energies of the two kinds of particles compare? (*b*) How do their average speeds compare?

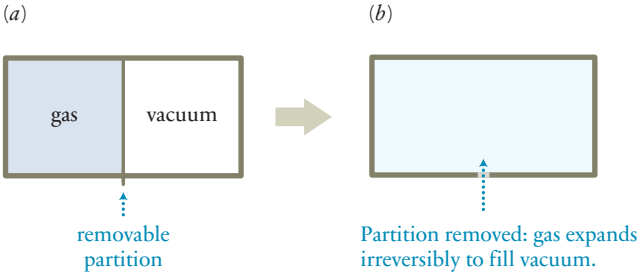
19.3 Equipartition of space

Collisions not only randomize the energy distribution in a gas, they also randomize spatial distributions. Consider, for example, a container divided into two compartments by a removable partition (**Figure 19.9**). In the left compartment is a gas, and the right compartment is empty. If we remove the partition, the gas diffuses to fill the entire volume. Likewise, if we put one type of gas on one side and another type of gas on the other side and then remove the partition, the two gases diffuse and mix. After some length of time has passed, both types of particles are evenly distributed throughout the container.

Both the diffusion of a single gas and the mixing of two different gases are irreversible processes. In Figure 19.9, for example, once the particles have expanded into the right side of the container, they do not spontaneously contract again into the left side.

Like dissipation of energy, this irreversible diffusion is due to the incoherent motion of particles at the atomic level. Collisions between the particles jumble up both

Figure 19.9 (*a*) A container has gas in one half and vacuum in the other, divided by a partition. (*b*) When the partition is removed, the gas expands irreversibly into the vacuum. After the expansion is complete, there is an equal likelihood of finding a gas particle anywhere in the container.



CONCEPTS

the speeds of the particles and the directions in which they move. Consequently the likelihood at any instant of finding a particle at any given position in the container is the same as at any other position or instant—there is no "preferred" position for the particles.

19.9 Suppose there are initially two indistinguishable particles in the left compartment of the container in Figure 19.9*a*. After the partition is removed, what is the probability at any given instant of finding one particle in each half of the container?

Checkpoint 19.9 suggests that we can use the statistical approach of Section 19.2 to determine the most likely locations of particles in a container. Consider, for example, a box that contains six indistinguishable particles. In order to count basic states, we need to "digitize" space just as we digitized energy in Section 19.2. So let's divide the container into four equal compartments as shown in Figure 19.10 (there are no real partitions).

The basic state of this system is determined by specifying how many particles are in each of the four compartments. There are many ways to characterize the system. Let's arbitrarily define the macrostate of the system to be the number of particles in the upper left compartment. (An alternative macrostate would be the number of particles in the lower right compartment, say, or the number in the lower half.) In Figure 19.10, all six particles are in the upper left compartment. The basic state is thus specified by the set of numbers (6, 0, 0, 0), representing the number of particles in each compartment, and the macrostate by the number 6 (for the number of particles in the upper left compartment). To determine the probability of having this distribution rather than any other, we need to determine the number of basic states and then the probability of each macrostate.

EXERCISE 19.4 Spatial macrostates and basic states

In Figure 19.10, how many basic states correspond to the macrostate in which (*a*) all six particles are in the top left compartment and (*b*) five particles are in the upper left compartment?

SOLUTION (*a*) One basic state because there is only one way to put all six particles in the upper left compartment (Figure 19.11*a*). ✔

(*b*) When five particles are in the upper left compartment, the last particle can be placed in any of the three other compartments. These three basic states are shown in Figure 19.11*b*. ✔

As you may have noticed, the problem of determining the distribution of six particles in four equal compartments of a container is identical to the energy distribution problem of Section 19.2, where we distributed six indistinguishable units of energy among four distinguishable "energy containers." Here we are distributing six indistinguishable "units of matter" among four distinguishable "spatial containers." The correspondence between the two problems is summarized in Table 19.1.

Figure 19.10

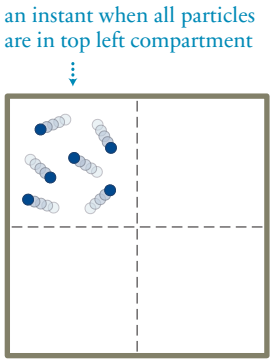


Figure 19.11

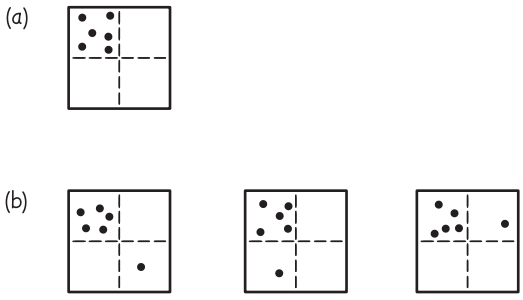
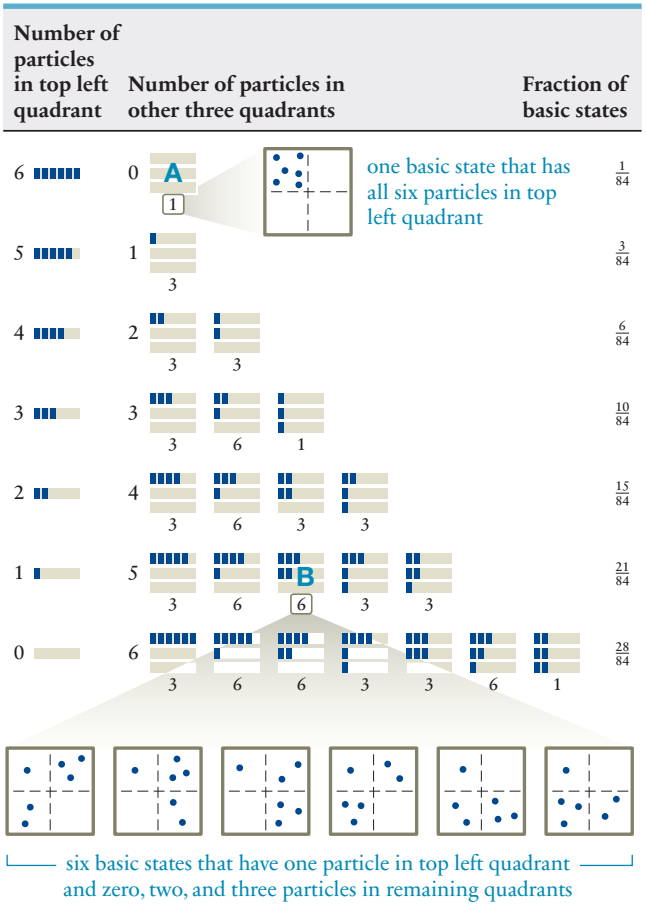


Table 19.1 Two problems, one idea

	Section 19.2	Section 19.3
phenomenon	distribution of energy	distribution of particles
indistinguishable items to be distributed	units of energy	particles
over distinguishable	pendulum and particles	compartments of container
macrostate defined by	energy units in pendulum	particles in top left compartment
basic state defined by	energy units in pendulum and each particle	particles in each compartment

CONCEPTS

Figure 19.12 Distribution of six particles in four equal compartments of a box.



Using the mapping from Table 19.1, we can adapt Figure 19.4 to our current discussion, as shown in Figure 19.12. The bar in the leftmost column, which represented the energy in the pendulum in Figure 19.4, now represents the number of particles in the top left compartment of the box. All the bars arranged in groups of three represent the number of particles in each of the other three compartments. The distribution labeled A thus represents the only basic state that corresponds to having all six particles in the top left compartment: The other three compartments are empty. The distribution labeled B represents six of the 21 basic states associated with the macrostate in which one particle is in the top left compartment. These basic states have three particles in any one of the other three compartments, two particles in another, and zero particles in the last remaining compartment.

19.10 (*a*) What is the probability of finding all six particles of Figure 19.10 in one compartment? (*b*) Of finding five particles in one compartment?

EXAMPLE 19.5 Particles in box

Use the data in Figure 19.12 to determine the average number of particles in the top left compartment of the four-compartment box shown there.

❶ **GETTING STARTED** The rightmost column in Figure 19.12 gives the fraction of basic states associated with each of the seven macrostates. This fraction represents the probability of finding the indicated number of particles in the top left compartment. For example, the probability of finding three particles there is $\frac{10}{84}$. In other words, if I observe the system for a long time interval, I'll see three particles in the top left compartment $\frac{10}{84}$th of the time.

❷ **DEVISE PLAN** To calculate the average number of particles in the top left compartment, I should take the sum over all macrostates of the product of the number of particles in that compartment (leftmost column in Figure 19.12) and the probability of finding that macrostate (rightmost column).

❸ **EXECUTE PLAN** Summing the products of the values in the leftmost and rightmost columns gives $6(\frac{1}{84}) + 5(\frac{3}{84}) + 4(\frac{6}{84}) + 3(\frac{10}{84}) + 2(\frac{15}{84}) + 1(\frac{21}{84}) + 0(\frac{28}{84}) = 1.5$. So, averaged over time, there are 1.5 particles in the top left compartment. ✔

❹ **EVALUATE RESULT** My answer makes sense because there are six particles in the box to be divided over four compartments: $6/4 = 1.5$.

Example 19.5 shows that even though the most probable macrostate has no particles in the upper left compartment, the average number of particles in that compartment is nonzero. Let us generalize our result by dividing the box into a large number M of equal compartments and increasing the number of particles N. On average each compartment tends to have its "share" N/M of the particles, yielding a uniform distribution of all the particles over space.

This tendency to distribute particles uniformly over space is called **equipartition of space** in analogy to the concept of equipartition of energy we encountered in Section 19.2. The diffusion of ink in water or smoke in air and the uniform distribution of gas particles in a container are all manifestations of this equipartitioning of space.

19.11 Suppose the partition in Figure 19.9 is designed to slide freely to the left and right. It is positioned in the center so that the container is divided into two equal compartments and clamped in place. Then $2N$ particles of a gas are added to the left compartment and N particles are added to the right compartment. When the clamp on the partition is removed, does the partition move to the right, move to the left, or stay where it is?

CONCEPTS

19.4 Evolution toward the most probable macrostate

In Sections 19.2 and 19.3, we assumed that the distribution of energy and the spatial distribution of particles are entirely random, like the throw of dice. However, there is an important difference between the throwing of dice and, say, the slowing down of a pendulum.

Each time you throw two dice, the combination you get is completely independent of the combinations you threw before. After throwing a sum of 7, for example, you can throw 12, 2, another 7, or any other sum. The dissipation of the energy in a swinging pendulum, on the other hand, takes many collisions. Because a single collision redistributes only a very small portion of the system's energy, the way the energy is distributed after any given collision is only slightly different from the way it was distributed just before the collision. In other words, the energy distribution in each post-collision macrostate is not independent of the energy distribution in the pre-collision macrostate.

19.12 In the system shown in Figures 19.3 and 19.4, suppose each collision transfers exactly one unit of energy from one colliding object to the other. (*a*) If the system is in the macrostate that has six energy units in the pendulum, which basic state(s) can the system be in after one collision? (*b*) If the system is in the macrostate that has five energy units in the pendulum, which *macrostates* can the system be in after one collision? (*c*) How many basic states do the macrostates in part *b* correspond to?

Checkpoint 19.12 shows that only neighboring macrostates are accessible to the system after each collision. So, if we start with all the energy in the pendulum, the system can only gradually evolve to the most probable macrostate (the one with all the energy in the gas particles). It takes a minimum of six collisions to reach the most probable macrostate. The number is likely to be much greater because, as you saw in doing the checkpoint, collisions can take place between two particles, leaving the energy in the pendulum unchanged, or a collision may *add* energy to the pendulum.

19.13 Suppose the system in Figure 19.4 is in the macrostate in which two energy units are in the pendulum and one collision takes place. (*a*) What are the accessible macrostates? (*b*) What is the probability for each of these macrostates?

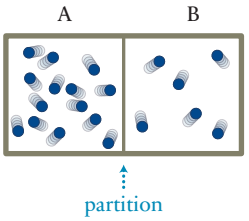
Note that the probability of reducing the energy in the pendulum is always higher than that of increasing the energy, and so the system gradually evolves to macrostates of increasing probability. Once the system is in the most probable macrostate, collisions continue to shuffle energy units around, but the system remains near this most probable macrostate. The system is therefore at equilibrium, and the most probable macrostate is the **equilibrium state.** If the pendulum and particles do temporarily occupy a macrostate that is less probable than the equilibrium state, then any subsequent collisions tend to move the system back toward equilibrium.

As we have seen, the probability of all the energy moving from the particles to the pendulum is vanishingly small, even for a gas of just 100 particles. This one-way flow of energy from pendulum to gas particles is the essence of *irreversibility* (see Section 5.3). Irreversibility is strongly connected to our sense of time. The direction of time is clear because everything around us is evolving irreversibly toward equilibrium. Now you can see that the underlying reason for irreversibility is statistics: Any system that is in a low-probability macrostate tends to evolve toward the macrostate that has the maximum number of basic states. Once the system has reached that equilibrium state, its probability of spontaneously returning to a macrostate that has a small number of basic states is vanishingly small.

To get a better sense of this evolution toward equilibrium, let's examine the particles in the closed box of **Figure 19.13**. The box contains 20 particles and is divided into compartments A and B. The partition prevents the particles from traveling from one compartment to the other but still allows particles from the two compartments to exchange energy with one another via collisions with the partition.

The energy of the system is constant—no energy can leak out because the system is closed. Suppose A contains 14 particles, B contains six particles, and the box contains ten units of energy. **Table 19.2** on the next page shows how these ten units can be distributed among the 20 particles. Let's define the macrostate by the energy units in compartment A. The number of basic states associated with each compartment is represented by the symbol Ω (the Greek capital omega) plus a subscript identifying the compartment. For example, there is only one basic state for compartment

Figure 19.13 Box containing 20 particles (14 on the left, 6 on the right) and 10 units of energy. The particles in the two compartments can exchange energy with one another through collisions with the partition.



CONCEPTS

Table 19.2 Number of basic states Ω associated with macrostates having 10 energy units distributed over 20 particles

E_A	E_B	Ω_A	Ω_B	$\Omega = \Omega_A\Omega_B$	ln Ω
Energy units in A	Energy units in B	Corresponding number of basic states for A	Corresponding number of basic states for B	Corresponding number of basic states for macrostate	
0	10	1	3003	3.00×10^3	8.01
1	9	14	2002	2.80×10^4	10.2
2	8	105	1287	1.35×10^5	11.8
3	7	560	792	4.44×10^5	13.0
4	6	2380	462	1.10×10^6	13.9
5	5	8568	252	2.16×10^6	14.6
6	4	27,132	126	3.42×10^6	15.0
7	3	77,520	56	4.34×10^6	15.3
8	2	203,490	21	4.27×10^6	15.3
9	1	497,420	6	2.98×10^6	14.9
10	0	1,144,066	1	1.14×10^6	14.0
Total		1,961,256	8008	$\Omega_{tot} = 2.00 \times 10^7$	

A ($\Omega_A = 1$) for macrostate $E_A = 0$. With nine energy units in A ($E_A = 9$) and the other one in B, there are 497,420 basic states for A ($\Omega_A = 497{,}420$) and six basic states for B ($\Omega_B = 6$).

19.14 How many ways are there to distribute two indistinguishable energy units among six distinguishable particles?

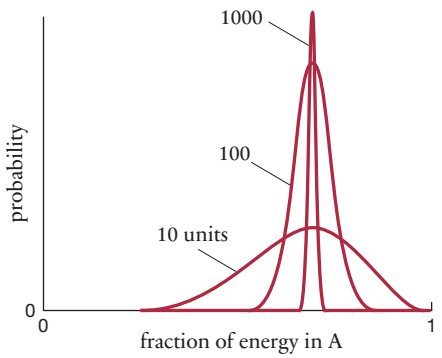
CONCEPTS

Note in Table 19.2 that as the number of energy units in each compartment of the box increases, the number of basic states in that compartment increases rapidly. Because the number of energy units in the box is constant, the number of basic states in one compartment decreases as the number of basic states in the other compartment increases. What then is the most probable macrostate for the box as a whole?

To answer this question, consider the macrostate $E_A = 1$ in Table 19.2. There are 14 ways to distribute this 1 energy unit in A (it can go into any one of the 14 particles of A), so there are 14 basic states associated with compartment A for $E_A = 1$. One energy unit in A leaves nine energy units to be distributed in B, and as you can see from the table, there are 2002 ways to distribute these nine units over the six particles in B. For each of the 14 basic states associated with compartment A, there are 2002 basic states associated with compartment B. The number of basic states associated with the macrostate $E_A = 1$ is thus $14 \times 2002 = 28{,}028$. More generally, for a given macrostate, the number Ω of basic states available to the system is obtained by multiplying Ω_A by Ω_B: $\Omega = \Omega_A\Omega_B$.

The probability of each macrostate is obtained by dividing Ω, the number of basic states associated with that macrostate, by Ω_{tot}, the number of basic states associated with all macrostates (2.00×10^7; see Table 19.2). The table shows you that this probability is greatest for the macrostate $E_A = 7$, as you would expect. Given that there are 14 particles in A and six in B, on average each particle has half an energy unit, and so the $E_A = 7$ macrostate corresponds to an equipartitioning of the energy. The curve labeled 10 units in **Figure 19.14** shows this probability as a function of the fraction of energy contained in A.

Figure 19.14 Probability of finding a given fraction of the system's energy in compartment A of the box in Figure 19.13. As the number of energy units increases from 10 to 1000, the probability distribution becomes narrower but remains centered about the mean energy.



EXAMPLE 19.6 Probability of macrostates

In Figure 19.13, after a very large number of particle-partition collisions have occurred, what is the probability of finding the system in (*a*) the macrostate $E_A = 1$ and (*b*) the macrostate $E_A = 7$?

1 GETTING STARTED Because all basic states are equally likely, the probability of finding the system in macrostate E_A is equal to the fraction Ω/Ω_{tot}, where Ω is the number of basic states of the system associated with the macrostate E_A and Ω_{tot} is the number of basic states associated with all macrostates (2.00×10^7; Table 19.2).

2 DEVISE PLAN To calculate the probability of a given macrostate E_A, I divide the value of Ω for that macrostate given in Table 19.2 by $\Omega_{tot} = 2.00 \times 10^7$.

3 EXECUTE PLAN (*a*) For $E_A = 1$, Table 19.2 tells me that $\Omega = 2.80 \times 10^4$. The probability of macrostate $E_A = 1$ is thus $(2.80 \times 10^4)/(2.00 \times 10^7) = 1.40 \times 10^{-3}$. ✔

(*b*) For macrostate $E_A = 7$, $\Omega = 4.34 \times 10^6$. So the probability of this macrostate occurring is $(4.34 \times 10^6)/(2.00 \times 10^7) = 2.17 \times 10^{-1}$. ✔

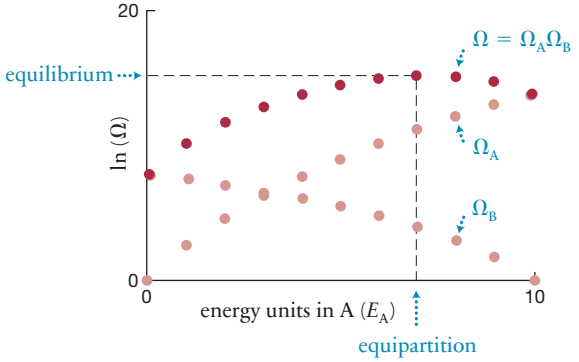
4 EVALUATE RESULT My result shows that the macrostate $E_A = 7$ is more than 150 times more probable than the macrostate $E_A = 1$. This makes sense because, as we saw earlier, the macrostate $E_A = 7$ is the equilibrium state for which there is an equipartition of energy.

If we increase the number of energy units in the box of Figure 19.13 to 100 or 1000, the number of basic states grows exponentially, and if we plot the probability of each macrostate as a function of the fraction of energy in A, we obtain the two curves labeled 100 and 1000 in Figure 19.14. Just as we saw in Figure 19.7, the most probable macrostate doesn't change, but the probability peaks much more narrowly around this state. In other words, the most probable macrostate—the equilibrium state—is now even more likely than any other macrostate.

Note that the number of basic states is very large, even with just ten energy units and 20 particles. In a box of volume 1 m^3 containing air at atmospheric pressure and room temperature, there are on the order of 10^{25} particles and 10^{20} energy units per particle, and so the number of basic states becomes unimaginably large—on the order of ten raised to the power 10^{21}! Because the number of basic states is so large, it is more convenient to work with the natural logarithm of that number. As you can see from the rightmost column in Table 19.2, the natural logarithm of the number of basic states is indeed much more manageable.

Figure 19.15 shows how the natural logarithms of Ω_A, Ω_B, and Ω vary with the number of energy units in compartment A in Figure 19.13. As you can see, the natural logarithm of the number of basic states changes much less rapidly than the number of basic states. Note that as E_A increases, the number of basic states Ω_A increases. As

Figure 19.15 Natural logarithm of the number of basic states for compartment A, for compartment B, and for the two compartments in Figure 19.13 combined. The number of basic states is maximal when the energy is equipartitioned (seven energy units in A).

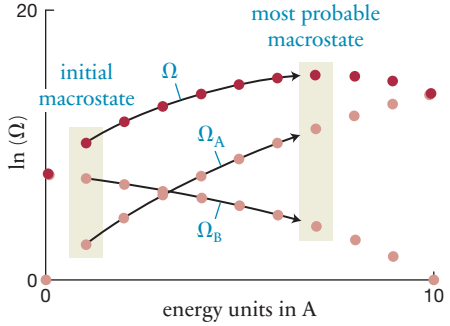


E_A increases, however, E_B decreases and so Ω_B decreases. The number of basic states Ω is maximum when $E_A = 7$ and $E_B = 3$, representing an equipartition of energy. The most probable macrostate (equilibrium) is achieved when there is equipartition of energy.

19.15 What is the average energy per particle in compartments A and B in Figure 19.13 (*a*) when there is one energy unit in A and (*b*) when the system is at equilibrium?

As you can see from Table 19.2, with $E_A = 1$ the number of basic states for the system (2.80×10^4) is more than 100 times smaller than it is at equilibrium ($E_A = 7$, $\Omega = 4.34 \times 10^6$). Collisions between the particles and the partition redistribute the energy among the particles, and over time the system evolves toward the most probable macrostate—the one for which each particle has its "fair share" of the system's energy (10/20 = 0.5 unit), just as you found in Checkpoint 19.15*b*. As the system evolves toward equilibrium from the $E_A = 1$ macrostate (**Figure 19.16**), the number of basic states in B *decreases*; that is, B

Figure 19.16 If the system evolves toward equilibrium from an initial macrostate with 1 energy unit in A, the number of basic states increases in A and decreases in B. The decrease for B is smaller than the increase for A and so the number of basic states still increases.



CONCEPTS

moves toward a macrostate of *lower* probability. However, the rate at which Ω_A increases is greater than the rate at which Ω_B decreases, and so Ω still increases.

EXAMPLE 19.7 Equilibrium

Suppose the system depicted in Figure 19.13 evolves toward equilibrium from the macrostate $E_A = 4$, $E_B = 6$. Determine the changes in (*a*) the number of energy units and (*b*) the number of basic states for compartment A, for compartment B, and for the system.

1 GETTING STARTED From Table 19.2 I see that the equilibrium macrostate (the one that has the greatest number of basic states associated with it) has $E_A = 7$ and $E_B = 3$. This is the macrostate toward which the system evolves.

2 DEVISE PLAN The change in the number of energy units in A is $\Delta E_A = E_{A,f} - E_{A,i}$. The change in the number of energy units in B is $\Delta E_B = E_{B,f} - E_{B,i}$. The change for this system is the sum of these two values. To determine the changes in the number of basic states, I obtain the number of basic states for the initial and final macrostates from Table 19.2 and compute the differences.

3 EXECUTE PLAN (*a*)

$$\Delta E_A = E_{A,f} - E_{A,i} = 7 - 4 = +3;$$
$$\Delta E_B = E_{B,f} - E_{B,i} = 3 - 6 = -3;$$
$$\Delta E = \Delta E_A + \Delta E_B = 3 - 3 = 0. \checkmark$$

(*b*) $\Delta\Omega_A = \Omega_{A,f} - \Omega_{A,i} = 77{,}520 - 2380 = 75{,}140$

$\Delta\Omega_B = \Omega_{B,f} - \Omega_{B,i} = 56 - 462 = -406$

$\Delta\Omega = \Omega_f - \Omega_i = (\Omega_{A,f})(\Omega_{B,f}) - (\Omega_{A,i})(\Omega_{B,i})$

$= (4.34 \times 10^6) - (1.10 \times 10^6) = 3.24 \times 10^6. \checkmark$

4 EVALUATE RESULT My answer to part *a* shows that the change in the number of energy units in the system is zero, as I expect because the system is closed. My result in *b* shows that even though Ω_B goes down, Ω_A increases significantly more, and so there is a gain in the number of basic states, as should happen when a system evolves toward equilibrium. Because Ω is the product of Ω_A and Ω_B, $\Delta\Omega$ is not simply the sum of $\Delta\Omega_A$ and $\Delta\Omega_B$. My result for $\Delta\Omega$, however, shows an increase in the number of basic states, as I expect.

The evolution toward equilibrium in Figure 19.13 (or in any other system) is a gradual process that seeks out the most probable macrostates through a very large number of random collisions. Once the system has reached equilibrium, the probability of returning to the macrostate $E_A = 1$ energy unit is very small—even for a system of just 20 particles. For a more realistic system, in which the number of particles is much greater, the odds favor equipartitioning so strongly that the probability of observing a system away from equilibrium is virtually zero unless the system is disturbed by outside influences.

The irreversible evolution of a system toward equilibrium can be quantified by the number of basic states Ω:

> **A closed system always evolves so as to maximize the number of basic states Ω. When this number has reached a maximum, the system is at equilibrium.**

This irreversible tendency of closed systems to maximize Ω is called the **second law of thermodynamics.***

Often we cannot directly determine the number of basic states Ω for a system. The second law of thermodynamics, however, allows us to reach conclusions about Ω in a closed system without having to consider any details at the atomic level. Consider, for example, a closed system consisting of a drop of ink placed in a beaker of water. We know from experience that the ink diffuses until it is uniformly distributed in the water (equipartition of space). In Section 19.3 we saw that the spreading out of particles increases the number of basic states—Ω for six particles in a container is much greater at equilibrium than when all the particles are clustered in a corner. For the ink, too, Ω increases as the ink spreads out, and so we can use our intuitive sense of the direction of time to reach conclusions about the number of basic states: If a process in a closed system runs only one way (the process is irreversible), the number of basic states of that system must be increasing. Conversely, the number of basic states of a closed system undergoing a reversible process remains constant.

CONCEPTS

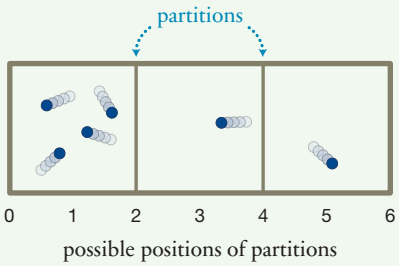
19.16 For which of these systems does Ω increase over time: (*a*) a thermos bottle containing warm water with an ice cube in it, (*b*) a cup of tea that is at the same temperature as the air in an insulated room, (*c*) a well-shaken bottle containing oil and water, (*d*) a seed germinating in soil inside a closed container, (*e*) the universe.

**Thermodynamics* is the branch of physics that deals with thermal energy and devices such as engines, subjects we discuss in the next two chapters. The second law of thermodynamics was first formulated (in a different form) in the context of this branch of physics. The formulation given here represents the modern and more general view of this law.

Self-quiz

1. Figure 19.4 shows that there are 15 basic states for four energy units distributed over three particles. List all 15.
2. Suppose a container holds three types of gas. The masses of the gas particles are m, $4m$, and $9m$. How do the average speeds of the particles compare?
3. In Figure 19.12 what is the probability of finding three particles in any compartment of the box?
4. The two partitions in the container shown in **Figure 19.17** can slide freely. Where will they be positioned when the space is equipartitioned?

Figure 19.17



5. A box with a fixed partition through which energy can be exchanged contains 50 particles, 40 on the left and 10 on the right. The system contains 250 energy units. As the system approaches equilibrium, how many energy units end up on the left? On the right?
6. If a closed system tends to change in a specific direction, such as ice melting in an insulated glass of water, what can you say about the number of basic states for this system?

Answers

1. 400, 040, 004, 310, 301, 031, 130, 013, 103, 220, 202, 022, 112, 121, 211.
2. According to the concept of equipartition of energy, the energy needs to be shared equally among all the gas particles, $\frac{1}{2}mv_1^2 = \frac{1}{2}(4m)v_2^2 = \frac{1}{2}(9m)v_3^2$, or $v_1^2 = 4v_2^2 = 9v_3^2$. Thus $v_1 = 2v_2 = 3v_3$.
3. Figure 19.12 shows that there are ten basic states with three particles in the top left compartment, regardless of the number of particles in the other compartments. When there are two particles in the top left compartment, there are six basic states with three particles in any other compartment. With one particle in the top left compartment, there are $6 + 3 = 9$ basic states with three particles in any other compartment. With no particles in the top left compartment, there are $6 + 3 = 9$ basic states with three particles in any other compartment. Therefore there are $10 + 6 + 9 + 9 = 34$ basic states with three particles in any compartment. The probability of finding three particles in any compartment is thus 34/84.
4. The partition shown at position 2 moves to position 4, and the partition shown at 4 moves to 5, so that there is 1 particle per unit length of the box.
5. Because at equilibrium the energy must be equipartitioned, the 250 units are evenly distributed across the 50 particles; with five units per particle. The 40 particles on the left have 200 units of energy, and the 10 particles on the right have 50 units.
6. According to the second law of thermodynamics, the system evolves to the macrostate that has the maximum number of basic states. Because the process is irreversible, the final macrostate must have a greater number of basic states than the initial macrostate.

19.5 Dependence of entropy on volume

In Sections 19.2 and 19.3, we saw that the equipartition of energy and space describes the most probable macrostate of any system—that is, the state that has the greatest number of basic states. The number of *atomic parameters*, such as the position and velocity of each particle, is intractably large for any real system, and so it is generally impossible to specify the basic state of the system. The macrostate, on the other hand, is readily specified by a few measurable *macroscopic parameters*, such as volume, pressure, or temperature.

To make the connection between macroscopic and atomic parameters, we consider a gas that consists of a very large number N of particles in a container of volume V. The particles occupy a negligible fraction of the volume V and interact with one another only during collisions that randomize both the energy distribution and the spatial distribution. All collisions between the particles and the walls of the container are elastic. A gas that satisfies these conditions is called an **ideal gas.** At high temperature and low pressure, an ideal gas is a good model for a real gas.

Let us now calculate how the number of spatial basic states depends on V and N. The system is closed so that the energy of the gas remains constant. Throughout this section we shall take this energy to be equipartitioned over the gas particles. At equilibrium each particle has an equal probability of occupying any location inside the container; the space is equipartitioned and each particle has its fair share of the volume V of the container. To count basic states, we divide the container into equal-sized compartments as we did in Section 19.3. If there are M equal-sized compartments and just one particle, that particle can be placed in any one of the M compartments. Thus there are M basic states for that one particle (**Figure 19.18**). A second particle can also be placed in any of the M compartments, and so, for two particles, there are M^2 basic states. Continuing this process for N particles, we see that Ω, the number of basic states, is*

$$\Omega = M^N \quad \text{(closed system)}. \tag{19.1}$$

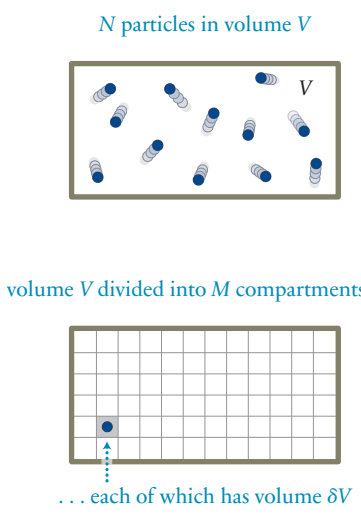


Figure 19.18 Each of the N particles inside the container can be placed in any one of the M equal-sized compartments.

EXERCISE 19.8 Number of basic states

Suppose ten distinguishable particles are equipartitioned in a container that is divided into 100 equal-sized compartments. (*a*) How many basic states are associated with this system? (*b*) What is the natural logarithm of this number of basic states?

SOLUTION (*a*) The number of basic states is $\Omega = M^N = 100^{10} = (10^2)^{10} = 10^{20}$, an unimaginably large number. ✔

(*b*) The natural logarithm of the number of basic states is $\ln \Omega = \ln(10^{20}) = 46$, a much more manageable number. ✔

19.17 If you decrease the size of the compartments in Exercise 19.8 by a factor of 10, what is the change (*a*) in the number of basic states and (*b*) in the natural logarithm of that number?

Exercise 19.8 and Checkpoint 19.17 illustrate that the natural logarithm of the number of basic states is much more manageable and less sensitive to a change in compartment size than the number itself. The natural logarithm of the number of basic states from Eq. 19.1 is

$$\ln \Omega = \ln(M^N) = N \ln M \quad \text{(closed system at equilibrium)}, \tag{19.2}$$

*We assume here that the particles are distinguishable from one another. In reality, the particles of an ideal gas are indistinguishable. In the case of indistinguishable particles and when $M \gg N \gg 1$, the number of basic states is found to be $M^N/N!$ instead of the result shown in Eq. 19.1. For simplicity, we shall limit the treatment in this section to systems of distinguishable particles.

QUANTITATIVE TOOLS

where we have used the identity $\ln(a^b) = b \ln a$. If the volume of the equal-sized compartments is δV, the number of compartments is $M = V/\delta V$, and so Eq. 19.2 becomes

$$\ln \Omega = N \ln\left(\frac{V}{\delta V}\right)$$
$$= N \ln V - N \ln \delta V \quad \text{(closed system at equilibrium)}, \qquad (19.3)$$

where we have used the identify $\ln(a/b) = \ln a - \ln b$.

Because the quantity $\ln \Omega$—the natural logarithm of the number of basic states—is so important, we define a quantity called **entropy:***

$$S \equiv \ln \Omega. \qquad (19.4)$$

Entropy is a unitless quantity that is a measure of the number of basic states in a system. The greater the number of basic states, the greater the entropy. Entropy allows us to quantify the spreading out of energy or particles over space.

As we saw in Section 19.4, a closed system always evolves so as to maximize the number of basic states. As the number of basic states increases, so does the entropy: $S_f > S_i$. Therefore,

$$\Delta S > 0 \quad \text{(closed system evolving toward equilibrium)}. \qquad (19.5)$$

At equilibrium, the number of basic states reaches its maximum value Ω_{max} and the entropy of the system no longer changes:

$$\Delta S = 0 \quad \text{(closed system at equilibrium)}. \qquad (19.6)$$

Equations 19.5 and 19.6 constitute the **entropy law,** a mathematical expression of the second law of thermodynamics. The entropy law implies that the entropy of a closed system never decreases (it can only increase or remain constant).

It is important to note that Eqs. 19.5 and 19.6 apply to only closed systems. For a system that is not closed, the entropy can increase, remain the same, or decrease. Back in Figure 19.16, for example, the particles in B transfer some of their energy to the particles in A, so B does not constitute a closed system. As we saw, Ω_B decreases and so the entropy of B decreases:

For systems that are not closed, the entropy can increase, decrease, or stay the same.

For a system of N gas particles equipartitioned over a volume V, we can obtain an expression for how entropy depends on volume by substituting Eq. 19.3 into Eq. 19.4:

$$S = N \ln V - N \ln \delta V \quad \text{(closed system at equilibrium)}. \qquad (19.7)$$

This expression describes a gas that is equipartitioned over a volume V—that is, at equilibrium. The value for S given by this expression thus represents the *maximum* value for the entropy of a gas of N particles in a volume V. If the gas is not at equilibrium, its entropy is smaller than that given by Eq. 19.7.

*Historically, entropy has been defined as $S \equiv k_B \ln \Omega$, where k_B is a constant that makes the entropy a much smaller number and gives it units of J/K (see Eq. 19.39). In this book, however, we shall use the definition given in Eq. 19.4. This definition is called a *statistical definition* of entropy.

The first term on the right in Eq. 19.7 shows that for a fixed number of particles N, the entropy (that is, the natural logarithm of the number of basic states) increases as we increase the volume V of the container. The second term on the right depends on the size of the compartments: The smaller the compartments, the smaller this term and therefore the greater the entropy. That the entropy depends on the volume δV of the compartments may seem disturbing—after all, the partitioning is completely arbitrary. In general, however, $\delta V \ll V$, and so the second term on the right in Eq. 19.7 is much smaller than the first one and can be ignored.

In general, we shall be concerned only with *changes* in entropy, in which case the second term in Eq. 19.7 drops out entirely. Consider, for example, the expansion of a gas made up of a fixed number of particles from an initial volume V_i to a final volume $V_f > V_i$. According to Eq. 19.7, the change in entropy is

$$\begin{aligned}\Delta S &= S_f - S_i = (N \ln V_f - N \ln \delta V) - (N \ln V_i - N \ln \delta V) \\ &= N \ln V_f - N \ln V_i = N \ln\left(\frac{V_f}{V_i}\right) \quad \text{(closed system at equilibrium)}.\end{aligned} \tag{19.8}$$

Because the second term on the right in Eq. 19.7 cancels out, the entropy change does not depend on the size of the compartments. Note that because $V_f > V_i$ the change in entropy is positive.

19.18 (*a*) What does a positive entropy change in a closed system imply about the change in the number of basic states? (*b*) How does Eq. 19.8 show that a gas expands if given more room? (*c*) How does it show that the gas does not contract spontaneously into a subvolume?

Let us next examine how entropy changes when a system of two interacting gases evolves to equilibrium. Consider, for example, two gases inside a container with a partition separating the gases into two compartments, one of volume V_A containing N_A particles and the other of volume V_B containing N_B particles (**Figure 19.19***a*).* If we let the partition move freely, the system evolves to the macrostate that has the maximum number of basic states. If the partition moves to the right, V_A increases, and so the number of basic states for gas A increases. As V_A increases, V_B decreases, and so the number of basic states for gas B decreases. Where does the partition come to rest? In other words, what value of V_A maximizes the number of basic states for the system as a whole?

In Section 19.4, we determined that the number of basic states for a system comprising two gases is the product of the number of basic states for the two separate gases:

$$\Omega = \Omega_A \Omega_B, \tag{19.9}$$

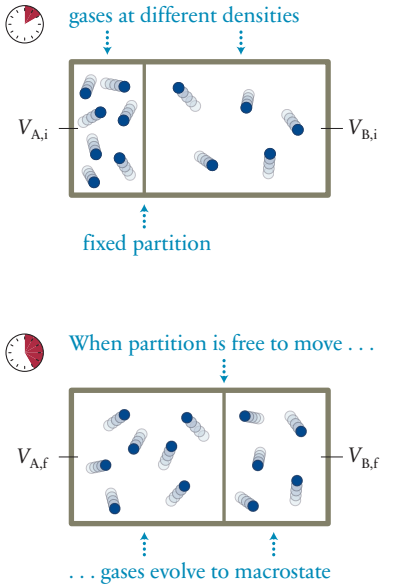
Using the definition of entropy (Eq. 19.4), we obtain

$$S = \ln(\Omega_A \Omega_B) = \ln \Omega_A + \ln \Omega_B = S_A + S_B, \tag{19.10}$$

where we have used the fact that $\ln(ab) = \ln a + \ln b$. Note that, unlike Ω, the natural logarithm of Ω and thus entropy are extensive properties:

The entropy of a combined system is the sum of the entropies of the individual systems.

Figure 19.19 (*a*) Two gases of unequal densities are placed on either side of a partition. (*b*) When the partition is free to move, the two gases evolve to the macrostate corresponding to the maximum number of basic states.



*We assume energy is equipartitioned over the system made up of both compartments and focus on only volume changes here. In the next section we shall examine the equipartitioning of energy.

QUANTITATIVE TOOLS

EXAMPLE 19.9 Partitioned gas

A partition divides a container into two equal compartments, A and B. Initially compartment A contains a gas of distinguishable particles and compartment B is empty. (*a*) What is the change in the entropy of the gas if the partition is removed and the gas is allowed to expand into the volume V of the container? (*b*) If the partition is reinserted after the gas has reached equilibrium in the volume V, what is the change in the entropy of the gas due to the reinsertion?

❶ **GETTING STARTED** When the partition is removed, the gas expands to fill the container, so the final volume occupied by the gas is twice the initial volume: $V_f = 2V_i$. Reinserting the partition divides the container into two equal compartments A and B again. The only difference is that now B is not empty; it contains half of the particles, and A contains the other half.

❷ **DEVISE PLAN** To obtain the change in entropy due to the removal of the partition, I substitute $V_f = 2V_i$ into Eq. 19.8. For part *b*, the entropy change due to the reinsertion is equal to the difference between the entropy of the gas after the partition is reinserted and its entropy before the partition is reinserted. To determine the entropy before the partition is reinserted, I use Eq. 19.7. Equation 19.10 tells me that after reinsertion the entropy of the gas is equal to the sum of the entropies in the two compartments. Because each compartment has half the volume of the container and contains half the gas particles, the entropies in the two compartments are equal to each other. To determine these entropies, I use Eq. 19.7 again.

❸ **EXECUTE PLAN** (*a*) Equation 19.8 gives $\Delta S = N \ln(2V_i/V_i) = N \ln 2$. ✔

(*b*) From Eq. 19.7, I obtain for the entropy before reinsertion of the partition: $S_i = N \ln V - N \ln \delta V$. Combining Eqs. 19.10 and 19.7, I obtain for the final entropy,

$$S_f = S_A + S_B = 2\left[\frac{N}{2}\ln\left(\frac{V/2}{\delta V}\right)\right] = N\ln\frac{V}{2} - N\ln\delta V$$
$$= N\ln V - N\ln 2 - N\ln\delta V.$$

The change in entropy is thus $\Delta S = S_f - S_i = -N \ln 2$. ✔

❹ **EVALUATE RESULT** My answer to part *a* shows that the removal of the partition and subsequent expansion of the gas result in a positive change in entropy. This is what I expect given that the system evolves toward equilibrium. My answer to part *b* is the negative of the answer to part *a*, which tells me that the change in entropy due to the combined removal and reinsertion of the partition is zero: $\Delta S = N \ln 2 - N \ln 2 = 0$. That is surprising because the gas has expanded into a larger volume. However, the particles in this problem are distinguishable, and in reality gas particles are indistinguishable. For distinguishable particles, the reinsertion reduces the number of basic states (a particle in compartment B can no longer be in compartment A, and vice versa). If the particles are indistinguishable, reinserting the partition causes no entropy change.

Assuming that the gases on each side of the partition in Figure 19.19 remain at equilibrium, we have from Eq. 19.7

$$S_A = N_A \ln V_A - N_A \ln \delta V \tag{19.11}$$

$$S_B = N_B \ln V_B - N_B \ln \delta V = N_B \ln(V - V_A) - N_B \ln \delta V, \tag{19.12}$$

where $V = V_A + V_B$ is the combined (constant) volume of the system. Increasing V_A thus increases S_A but decreases S_B. **Figure 19.20** shows how the entropies of the two compartments and the entropy of the combined system depend on V_A. When the number of basic states, and thus the entropy in Eq. 19.10, are maximum, Eq. 19.6 applies, so the partition comes to rest and the combined system is at equilibrium. We therefore would like to determine for what value of V_A the system's entropy S is maximized.

At the maximum value of S, the slope of the S curve in Figure 19.20 is zero. This slope is given by the derivative of S with respect to V_A, and so at equilibrium,

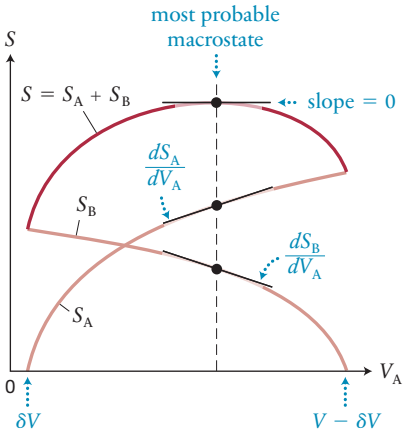
$$\frac{dS}{dV_A} = \frac{dS_A}{dV_A} + \frac{dS_B}{dV_A} = 0 \quad \text{(equilibrium)} \tag{19.13}$$

or

$$\frac{dS_A}{dV_A} = -\frac{dS_B}{dV_A} \quad \text{(equilibrium)}. \tag{19.14}$$

The quantities dS_A/dV_A and dS_B/dV_A represent the slopes of the S_A and S_B curves in Figure 19.20. Equation 19.14 thus tells us that these two curves must have opposite slopes at the equilibrium value of V_A.

Figure 19.20 As the partition in Figure 19.19 moves to the right and V_A increases, S_A increases and S_B decreases until the entropy of the system is maximum. At this value of V_A, the partition comes to rest, and the system is at equilibrium.



We can now show that the condition expressed in Eq. 19.14 leads to equipartitioning of the space by calculating the derivatives. Because V is constant,

$$\frac{dV_B}{dV_A} = \frac{d}{dV_A}(V - V_A) = -1, \tag{19.15}$$

we have $dV_A = -dV_B$ and so can rewrite Eq. 19.14 in the form

$$\frac{dS_A}{dV_A} = -\frac{dS_B}{(-dV_B)} = \frac{dS_B}{dV_B} \quad \text{(equilibrium)}. \tag{19.16}$$

Equation 19.16 tells us that at equilibrium the quantity dS/dV evaluated on either side of the partition is the same.

Taking the derivative of Eq. 19.7 with respect to V, we get

$$\frac{dS}{dV} = \frac{d}{dV}(N \ln V - N \ln \delta V) = \frac{d}{dV}(N \ln V) \tag{19.17}$$

because $N \ln \delta V$ does not depend on V (N is fixed and δV is a constant). Recalling that

$$\frac{d}{dx}\ln x = \frac{1}{x}, \tag{19.18}$$

we get

$$\frac{dS}{dV} = \frac{N}{V}. \tag{19.19}$$

The quantity N/V represents the number of particles per unit volume and is the number density of the gas (Eq. 1.3). Substituting Eq. 19.19 into Eq. 19.16, we obtain

$$\frac{N_A}{V_A} = \frac{N_B}{V_B} \quad \text{(equilibrium)}, \tag{19.20}$$

which is exactly the requirement for the equipartition of space: At equilibrium the number densities are the same on both sides of the partition, and so each particle has the same amount of space available to it.

EXAMPLE 19.10 Sliding partition

The box in Figure 19.19 contains seven gas particles in compartment A and five in compartment B, and the partition separating the compartments is free to move. Let the volume of the box be V and the initial ratio of the compartment volumes be $V_{B,i}/V_{A,i} = 3$. What is the change in the entropy of this gas as it evolves from this macrostate to equilibrium?

❶ **GETTING STARTED** Once the equilibrium state is reached, the number densities are equal on the two sides of the partition and the change in the entropy of the gas—which is my system—is the sum of the entropy changes in the two compartments.

❷ **DEVISE PLAN** Equation 19.20 gives me the volume ratio at equilibrium, which I call the final volume ratio. Knowing the initial and final ratios, I can use Eq. 19.8 to determine the entropy change in each compartment and then use Eq. 19.10 to determine the entropy change of the system.

❸ **EXECUTE PLAN** From the given initial volume ratio, I know that $V_{B,i} = 3V_{A,i}$. The volume of the box is the sum of the volumes of the compartments: $V = V_{A,i} + V_{B,i} = V_{A,i} + 3V_{A,i} = 4V_{A,i}$, and so $V_{A,i} = V/4$ and $V_{B,i} = 3V/4$. From Eq. 19.20 the ratio of the final volumes is equal to the ratio of the numbers of particles. I know that the number of particles in each compartment cannot change, so I can say that $V_{B,f}/V_{A,f} = N_B/N_A = 5/7$, which means $V_{B,f} = 5V_{A,f}/7$.

After the partition moves, the sum of the volumes is still V, and so I know $V = V_{A,f} + V_{B,f} = V_{A,f} + 5V_{A,f}/7 = 12V_{A,f}/7$.

QUANTITATIVE TOOLS

Thus $V_{A,f} = 7V/12$ and $V_{B,f} = 5V/12$. Substituting these volumes into Eq. 19.8 gives

$$\Delta S_A = N_A \ln \frac{V_{A,f}}{V_{A,i}} = 7 \ln \frac{7V/12}{V/4} = 7 \ln \tfrac{7}{3} = 5.93$$

$$\Delta S_B = N_B \ln \frac{V_{B,f}}{V_{B,i}} = 5 \ln \frac{5V/12}{3V/4} = 5 \ln \tfrac{5}{9} = -2.94.$$

Because the entropy of the system is the sum of the entropies of the particles in the two compartments (Eq. 19.10), I have

$$\Delta S = \Delta S_A + \Delta S_B = 5.93 - 2.94 = 2.99. \checkmark$$

4 EVALUATE RESULT The volume of compartment A increases and the volume of compartment B decreases, so I expect the entropy of the particles in A to increase and the entropy of the particles in B to decrease, as my results show. Because the magnitude of the change in entropy is greater for A than for B, the entropy change of the system is positive, as it should be for a closed system evolving toward equilibrium.

19.19 Explain why the changes in entropy in Example 19.10 do not depend on the volume V of the box.

19.6 Dependence of entropy on energy

Let us now turn our attention to how entropy depends on the energy of a system. For simplicity we restrict ourselves to a *monatomic* ideal gas (that is, a gas made up of atoms for which we can ignore any internal structure), so that none of the atoms' kinetic energy can be converted to internal energy. The incoherent kinetic energy of all the atoms makes up the thermal energy of the gas. If the center of mass of the gas is at rest, the energy of the system is entirely in the form of thermal energy.

Let us now calculate how the number of basic states depends on the thermal energy E_{th} and the number of atoms N of a closed system. Throughout this section we shall take the atoms to be distributed uniformly over the (constant) volume V of the system. At equilibrium each atom has its fair share of the thermal energy E_{th} of the system. A single atom moving at speed v contributes an amount of energy $\frac{1}{2}mv^2$ to E_{th}. Because the atom is free to move in any direction, many velocities correspond to a given speed v (**Figure 19.21**) and thus to a given share of the thermal energy E_{th}. A basic state is specified by the components v_x, v_y, and v_z of all N atoms. To count the basic states, we can use the procedure from Section 19.5: We divide the available "velocity space" in Figure 19.21 into small compartments and count the number of ways we can distribute N atoms among these compartments.

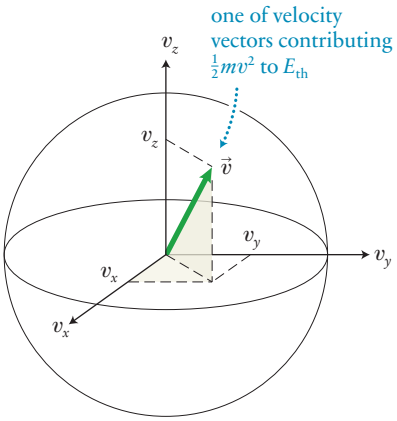
What is the available velocity space? To determine this velocity space, we introduce the **root-mean-square speed,*** or *rms speed* for short:

$$v_{rms} \equiv \sqrt{(v^2)_{av}}, \tag{19.21}$$

where $(v^2)_{av}$ is the average of the squares of the speeds of all the atoms in a volume of gas. Applying the Pythagorean theorem to the two shaded triangles in Figure 19.21, we have

$$v^2 = v_x^2 + v_y^2 + v_z^2. \tag{19.22}$$

Figure 19.21 Velocity space determined by the components of velocity. All velocity vectors ending on the surface of the sphere correspond to the same speed and therefore the same kinetic energy.



*This value is obtained by taking the square *root* of the *mean* (average) of the *squares* of the speeds.

If we average over all atoms, this expression becomes

$$v_{\text{rms}}^2 = (v^2)_{\text{av}} = (v_x^2)_{\text{av}} + (v_y^2)_{\text{av}} + (v_z^2)_{\text{av}}. \tag{19.23}$$

Given that the gas is the same in all directions, there is no difference between the x, y, and z directions, and so

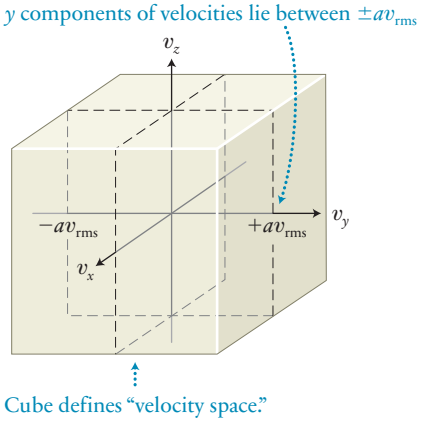
$$(v_x^2)_{\text{av}} = (v_y^2)_{\text{av}} = (v_z^2)_{\text{av}} \tag{19.24}$$

or

$$v_{\text{rms}}^2 = 3(v_x^2)_{\text{av}}. \tag{19.25}$$

19.20 Five atoms moving along the x axis have x components of velocities $+2.1$ m/s, -3.2 m/s, $+4.5$ m/s, -0.3 m/s, and -3.1 m/s. Determine (*a*) the x component of the average velocity, (*b*) the average speed, and (*c*) the rms speed.

Figure 19.22 Calculating the volume $\mathcal{V}$ of the velocity space.



To return to determining the volume of the velocity space, let us now assume that the components of the velocities of the atoms are limited to a range of a few times the rms speed. That is, let us assume that

$$-av_{\text{rms}} \le v_x \le av_{\text{rms}}, \tag{19.26}$$

and likewise for the y and z components, and where a is a number between, say, 1 and 10. This may seem like a totally arbitrary restriction, but you will see in a bit that the end result is not affected by this restriction. With this simplifying assumption the velocities of all the atoms in our system lie in a cube having sides of length $2av_{\text{rms}}$ (**Figure 19.22**). The "volume" $\mathcal{V}$ of the velocity space of the atoms in a gas-filled container is then $\mathcal{V} = (2av_{\text{rms}})^3$.

We can relate this volume to the thermal energy of the gas because we know that the average kinetic energy K_{av} of one atom is equal to the thermal energy of the gas divided by the number N of gas atoms in the container:

$$K_{\text{av}} \equiv \tfrac{1}{2}m(v^2)_{\text{av}} = \tfrac{1}{2}mv_{\text{rms}}^2 = \frac{E_{\text{th}}}{N} \quad \text{(monatomic ideal gas)}, \tag{19.27}$$

where we have used the definition of rms speed (Eq. 19.21). Now we can rewrite the rms speed as

$$v_{\text{rms}} = \left(\frac{2E_{\text{th}}}{mN}\right)^{1/2}, \tag{19.28}$$

and so we see that the volume $\mathcal{V}$ of the velocity space is proportional to $(E_{\text{th}})^{3/2}$:

$$\mathcal{V} = (2av_{\text{rms}})^3 = \left[2a\left(\frac{2E_{\text{th}}}{mN}\right)^{1/2}\right]^3 = \left(\frac{8a^2}{mN}\right)^{3/2}(E_{\text{th}})^{3/2} = b(E_{\text{th}})^{3/2}, \tag{19.29}$$

where $b = (8a^2/mN)^{3/2}$ is a constant.

19.21 Work out the algebra between the third and fourth terms in Eq. 19.29.

In analogy to what we did in Section 19.5, we now divide this volume $\mathcal{V}$ of the velocity space into smaller velocity-space compartments, each of volume $\delta\mathcal{V}$. The number of velocity-space compartments is

$$M_v = \frac{\mathcal{V}}{\delta\mathcal{V}} = \frac{b}{\delta\mathcal{V}}(E_{\text{th}})^{3/2}, \tag{19.30}$$

QUANTITATIVE TOOLS

and the number of ways we can distribute N distinguishable atoms in these M_v compartments is (see Eq. 19.1):

$$\Omega_v = (M_v)^N = \left[\frac{b}{\delta\mathcal{V}}(E_{\text{th}})^{3/2}\right]^N. \tag{19.31}$$

For a monatomic ideal gas in which the energy is equipartitioned, the entropy is maximum and given by

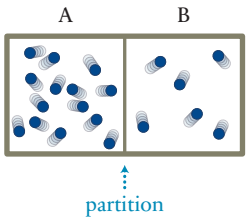
$$S \equiv \ln \Omega_v = \tfrac{3}{2}N \ln E_{\text{th}} - N \ln \frac{\delta\mathcal{V}}{b} \quad \text{(monatomic ideal gas at equilibrium).} \tag{19.32}$$

The first term to the right of the equals sign shows that for a fixed number of atoms N, the entropy increases as we increase the thermal energy E_{th} of the gas. The factors $\delta\mathcal{V}$ and b in Eq. 19.32 tell us that the entropy depends on the size of the velocity-space compartments in which the atoms are distributed and on the restriction we imposed in Eq. 19.26. However, the term that contains these factors drops out when we consider changes in entropy:

$$\Delta S = S_{\text{f}} - S_{\text{i}} = \tfrac{3}{2}N \ln E_{\text{th,f}} - \tfrac{3}{2}N \ln E_{\text{th,i}} = \tfrac{3}{2}N \ln\left(\frac{E_{\text{th,f}}}{E_{\text{th,i}}}\right). \tag{19.33}$$

Now we can examine how entropy changes when a system of two interacting gases initially having different amounts of thermal energy evolves toward a system-wide equilibrium state. Consider a closed system of two gases that can exchange energy with each other, as illustrated in **Figure 19.23**. The volume of each compartment and the number of atoms in each are fixed. In Section 19.4 we saw that such a system tends to evolve toward an equipartitioning of the system's thermal energy. If thermal energy is exchanged across the partition, one gas gains thermal energy and the other one loses thermal energy. Because the entropy of the gas depends on its thermal energy, the exchange of thermal energy causes the entropy to increase on one side of the partition and decrease on the other side. **Figure 19.24** shows how the entropies of the two gases and the entropy of the combined system depend on the thermal energy $E_{\text{th,A}}$ of gas A. When the number of basic states, and thus the entropy, are maximum, Eq. 19.6 applies and the exchange of thermal energy stops and the system is said to be in **thermal equilibrium.** We therefore would like to determine for what value of $E_{\text{th,A}}$ the system's entropy $S = S_{\text{A}} + S_{\text{B}}$ is maximized.

Figure 19.23 Container holding two gases separated by a partition. The gases can exchange energy with each other through collisions with the partition.



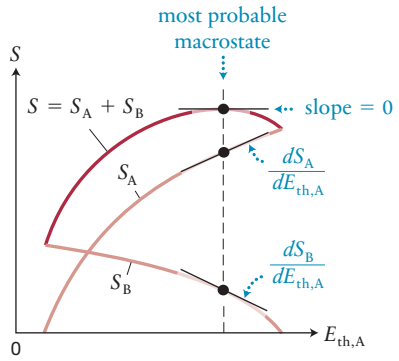
At the maximum entropy level, the slope of curve S in Figure 19.24 is zero, and so

$$\frac{dS}{dE_{\text{th,A}}} = \frac{dS_{\text{A}}}{dE_{\text{th,A}}} + \frac{dS_{\text{B}}}{dE_{\text{th,A}}} = 0 \quad \text{(closed system at thermal equilibrium)} \tag{19.34}$$

or

$$\frac{dS_{\text{A}}}{dE_{\text{th,A}}} = -\frac{dS_{\text{B}}}{dE_{\text{th,A}}} \quad \text{(closed system at thermal equilibrium).} \tag{19.35}$$

Figure 19.24 How the entropies of the gases in Figure 19.23 depend on the thermal energy $E_{\text{th,A}}$ of gas A.



The quantities $dS_{\text{A}}/dE_{\text{th,A}}$ and $dS_{\text{B}}/dE_{\text{th,A}}$ represent the slopes of curves S_{A} and S_{B} in Figure 19.24. Eq. 19.35 thus tells us that these two curves must have opposite slopes at thermal equilibrium. Because E_{th} is constant,

$$\frac{dE_{\text{th,B}}}{dE_{\text{th,A}}} = \frac{d}{dE_{\text{th,A}}}[E_{\text{th}} - E_{\text{th,A}}] = -1, \tag{19.36}$$

we have $dE_{\text{th,A}} = -dE_{\text{th,B}}$ and so Eq. 19.35 becomes

$$\frac{dS_A}{dE_{\text{th,A}}} = -\frac{dS_B}{d(-E_{\text{th,B}})} = \frac{dS_B}{dE_{\text{th,B}}} \quad \text{(closed system at thermal equilibrium).} \tag{19.37}$$

As you have probably noticed, there is a strong parallel between the quantity dS/dE_{th} in the above derivation and the quantity dS/dV we discussed in the preceding section (compare, also, Figures 19.20 and 19.24). In Eq. 19.19 we saw that the quantity dS/dV represents the number density of the gas and that at equilibrium the number density tends to be the same in all compartments of the system (Eq. 19.20).

To discover what physical concept dS/dE_{th} corresponds to, answer the following checkpoint.



19.22 (*a*) Using Figure 19.24, compare the kinetic energy of an atom in gas A with that of an atom in gas B in the thermal equilibrium state (that is, on the dashed line) and in a macrostate to the left of the dashed line. (*b*) In each of these two macrostates, which gas is "hotter"? (*c*) For the same two macrostates, compare the quantities $dS_A/dE_{\text{th,A}}$ and $dS_B/dE_{\text{th,B}}$.

As Checkpoint 19.22 suggests, the quantity dS/dE_{th} is smaller for a hot gas than for a cooler gas. We therefore define this quantity to be inversely proportional to a quantity called the **absolute temperature** T:

$$\frac{1}{k_B T} \equiv \frac{dS}{dE_{\text{th}}}. \tag{19.38}$$

The unit for absolute temperature, the **kelvin** (K), is one of the SI base units. In Section 20.2 we'll see how this unit is defined. The proportionality constant k_B, called the **Boltzmann constant,** is

$$k_B = 1.381 \times 10^{-23}\ \text{J/K}. \tag{19.39}$$

The constant k_B had been used to reconcile the units for absolute temperature with common temperature measurements (which we shall discuss in Chapter 20), but it is now defined as exactly 1.380649×10^{-23} J/K. For now, consider T to be a quantitative measure of the temperature obtained by adding 273.15 to the reading on a thermometer calibrated in degrees Celsius.

With this definition of absolute temperature, we see that thermal equilibrium (Eq. 19.37) implies that the absolute temperatures of the two gases are the same:

$$T_A = T_B \quad \text{(thermal equilibrium).} \tag{19.40}$$

We can obtain another consequence of the thermal equilibrium condition by taking the derivative of Eq. 19.32 with respect to E_{th}:

$$\frac{dS}{dE_{\text{th}}} = \frac{d}{dE_{\text{th}}}\left[\tfrac{3}{2}N \ln E_{\text{th}} - N \ln \frac{\delta \mathcal{V}}{b}\right] = \frac{d}{dE_{\text{th}}}\left[\tfrac{3}{2}N \ln E_{\text{th}}\right] = \tfrac{3}{2}\frac{N}{E_{\text{th}}}, \tag{19.41}$$

where the derivative of the term $N \ln(\delta \mathcal{V}/b)$ is zero because N, $\delta \mathcal{V}$, and b are constant. Substituting Eq. 19.41 into Eq. 19.37 and inverting the result yield

$$\frac{E_{\text{th,A}}}{N_A} = \frac{E_{\text{th,B}}}{N_B} \quad \text{(thermal equilibrium),} \tag{19.42}$$

which is exactly the requirement for equipartition of energy: At thermal equilibrium the amount of thermal energy per atom is the same on either side of the partition.

19.23 Suppose 14 atoms are in compartment A and six atoms are in compartment B, as shown in Figure 19.23. When thermal equilibrium energy is reached, what percentage of the thermal energy do you expect to find in each compartment?

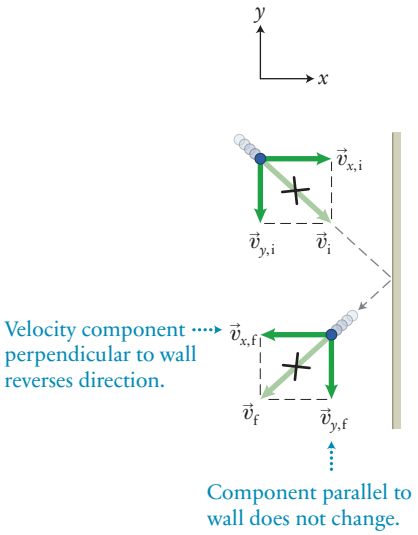
19.7 Properties of a monatomic ideal gas

Let us next connect some of the atomic parameters that describe an ideal gas to its macroscopic parameters. Consider a monatomic ideal gas that consists of a very large number N of indistinguishable atoms of mass m inside a closed rigid container of volume V. The collisions between the atoms and between the atoms and the walls of the container are all elastic. Figure 19.25 shows one atom-wall collision. Because the collision is elastic, the speed v of the atom does not change; the collision simply reverses the direction of the velocity component v_x perpendicular to the wall, leaving the velocity component v_y parallel to the wall unchanged.

What is true for the atom in Figure 19.25 is true for all the other gas atoms. Therefore the forces exerted by the wall and by the atoms are perpendicular to the wall. The collisions between the wall and the gas atoms are the origin of the pressure in the gas.

19.24 Let the system comprising the atoms and wall in Figure 19.25 be isolated. (*a*) What is the direction of the atom's change in momentum? (*b*) What is the direction of the change in momentum of the wall? (*c*) How are these momentum changes related to the forces that wall and atom exert on each other?

Figure 19.25 An atom collides elastically with a wall, reversing the perpendicular component of the incoming velocity. Its speed remains unchanged.



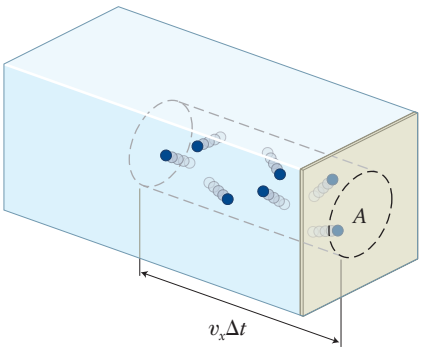
As we saw in Section 18.5, the pressure P in a gas is equal to the magnitude of the force $\vec{F}^{\,c}_{gA}$ exerted by the gas on a surface per unit area. Because the force exerted by the gas on the surface and the force exerted by the surface on the gas form a interaction pair, we have $F^c_{gA} = F^c_{Ag}$. Over a section of wall of area A, Eq. 18.1 then yields

$$P \equiv \frac{F^c_{gA}}{A} = \frac{F^c_{Ag}}{A}. \tag{19.43}$$

The force exerted by the wall on the gas is equal to the rate of change in the momentum of the gas (see Section 8.7), if we ignore other forces, and our task is to evaluate the magnitude of this rate of change. To simplify matters, we initially assume that the x component of the velocities of all the atoms in the gas have the same magnitude $|v_x|$. (We shall see shortly that our result doesn't depend on this oversimplification.) Let's begin by evaluating the change in momentum of a single gas atom when it collides with the wall. Each collision like the one in Figure 19.25 changes the sign of the x component of the atom's velocity, and so the magnitude of the atom's change in momentum is $2m|v_x|$.

To calculate the rate at which the magnitude of the momentum of the gas changes, we need to determine how many collisions occur within a finite time interval Δt. In order for an atom, for which the x component of its velocity is v_x, to collide with the wall in a time interval Δt, the atom has to be within a distance $|v_x|\Delta t$ of the wall. So let us consider all the atoms in a cylinder of cross-sectional area A and length $|v_x|\Delta t$ (Figure 19.26). At equilibrium, the atoms are uniformly distributed over the volume of the container, and so the number of atoms per unit volume is N/V. The volume of

Figure 19.26 On average, one half of the gas atoms contained in the cylinder strike the wall during a time interval Δt.



the cylinder is $A|v_x|\Delta t$, and so the number of atoms in the cylinder is $N_{\text{cyl}} = (N/V)A|v_x|\Delta t$. At any given instant, there should be as many atoms moving in the $+x$ direction as there are atoms moving in the $-x$ direction, and so half of the atoms in the cylinder in Figure 19.26 collide with the wall in the time interval Δt. The magnitude of the change in momentum for all gas atoms is thus equal to the number of collisions $N_{\text{cyl}}/2$ times the magnitude of the change in momentum $2m|v_x|$ for a single collision:

$$|\Delta p| = \tfrac{1}{2}\frac{N}{V}(A|v_x|\Delta t)(2m|v_x|) = \frac{NAmv_x^2}{V}\Delta t. \tag{19.44}$$

To calculate the magnitude of the rate of change of the momentum of the gas, we divide the left and right sides of the equation by Δt:

$$\frac{|\Delta p|}{\Delta t} = \frac{NAmv_x^2}{V}. \tag{19.45}$$

The quantity on the left, the magnitude of the rate of change of the momentum of the gas, is equal to the magnitude of the force exerted by the wall on the gas in the cylinder in Figure 19.26, and so, substituting Eq. 19.45 into Eq. 19.43, we obtain for the pressure in the gas

$$P = \frac{|\Delta p|/\Delta t}{A} = \frac{Nmv_x^2}{V}. \tag{19.46}$$

In deriving Eq. 19.46 we used the unlikely assumption that the x components of the velocities of all atoms have the same magnitude $|v_x|$. When we use this assumption, each atom contributes a force $(Amv_x^2)/V$, and the force in Eq. 19.45 is the sum of N such forces. If the atoms move at different speeds, we should replace v_x^2 in Eq. 19.46 by the average value $(v_x^2)_{\text{av}}$. (We can do this because N times the average value of a quantity is equal to the sum of N actual values.)

When we use the relationship between $(v_x^2)_{\text{av}}$ and v_{rms}^2 expressed in Eq. 19.25, $v_{\text{rms}}^2 = 3(v_x^2)_{\text{av}}$, Eq. 19.46 becomes

$$P = \frac{Nm\frac{1}{3}v_{\text{rms}}^2}{V} = \frac{2}{3}\frac{N(\frac{1}{2}mv_{\text{rms}}^2)}{V} \tag{19.47}$$

or, substituting $N(\frac{1}{2}mv_{\text{rms}}^2) = E_{\text{th}}$ (see Eq. 19.27),

$$P = \frac{2}{3}\frac{E_{\text{th}}}{V} \quad \text{(monatomic ideal gas)}. \tag{19.48}$$

Equation 19.47 shows that the pressure in a monatomic ideal gas is determined by two atomic parameters: the number density N/V and the average kinetic energy $\frac{1}{2}mv_{\text{rms}}^2$ of an atom.

We can obtain an expression for the thermal energy of a monatomic ideal gas by substituting Eq. 19.41, $dS/dE_{\text{th}} = \frac{3}{2}/E_{\text{th}}$, into Eq. 19.38, $1/k_{\text{B}}T = dS/dE_{\text{th}}$. Doing this we obtain

$$\frac{1}{k_{\text{B}}T} = \frac{3}{2}\frac{N}{E_{\text{th}}} \tag{19.49}$$

or

$$E_{\text{th}} = \tfrac{3}{2}Nk_{\text{B}}T \quad \text{(monatomic ideal gas)}. \tag{19.50}$$

QUANTITATIVE TOOLS

Substituting Eq. 19.50 into Eq. 19.48, we obtain

$$P = \frac{N}{V} k_B T \quad \text{(ideal gas system at equilibrium)}. \tag{19.51}$$

This relationship among pressure, volume, temperature, and number of atoms is called the **ideal gas law.** We obtained this result among macroscopic parameters (pressure, volume, temperature) by considering what happens at the atomic scale for a monatomic ideal gas. Experiments show that, at low pressure and high temperature, the ideal gas law describes the behavior of *any* gas at equilibrium, not just monatomic gases.

Because the thermal energy of a gas is the sum of the average kinetic energies of its atoms, $E_{th} = NK_{av}$ (Eq. 19.27), we can use the Eq. 19.50 expression for E_{th} and write

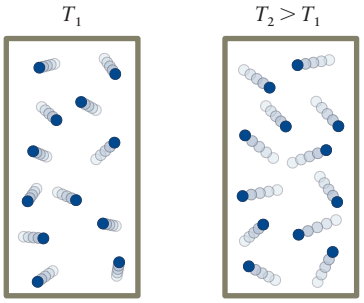
$$K_{av} = \tfrac{1}{2} m v_{rms}^2 = \tfrac{3}{2} k_B T. \tag{19.52}$$

Equation 19.52 allows us to obtain an expression for the rms speed of an atom in a monatomic ideal gas:

$$v_{rms} = \sqrt{\frac{3k_B T}{m}}. \tag{19.53}$$

The rms speed of a monatomic ideal gas atom is inversely proportional to the square root of the mass of the atom and proportional to the square root of the temperature of the gas. If the temperature of a gas increases, the rms speed of the atoms goes up, too (**Figure 19.27**).

Figure 19.27 When the temperature of a monatomic ideal gas in a fixed volume is increased, the rms speed of the gas atoms increases.



19.25 Use the ideal gas law to express the equilibrium condition in Eq. 19.20 in terms of pressure.

EXAMPLE 19.11 Helium speed

A sample of helium gas containing 6.02×10^{23} atoms at atmospheric pressure (1.01×10^5 Pa) occupies a volume of 22.4×10^{-3} m^3. The mass of a helium atom is 6.646×10^{-27} kg. What is the rms speed of the helium atoms?

❶ **GETTING STARTED** I can treat the helium gas as a monatomic ideal gas. The rms speed of the atoms is determined by the temperature of the sample, which is not given.

❷ **DEVISE PLAN** Equation 19.53 gives the rms speed of atoms in a monatomic ideal gas. I am given the mass m of one helium atom, but not the temperature T. To determine T I can use the ideal gas law (Eq. 19.51).

❸ **EXECUTE PLAN** From Eq. 19.51, I obtain $T = PV/(Nk_B)$. Substituting this expression for the temperature into Eq. 19.53 yields

$$v_{rms} = \sqrt{\frac{3k_B PV/(Nk_B)}{m}} = \sqrt{\frac{3PV}{mN}} = 1.30 \times 10^3 \text{ m/s}. \checkmark$$

❹ **EVALUATE RESULT** Those helium atoms sure are zipping around fast! At 1.30×10^3 m/s, however, they are moving about ten times more slowly than the escape speed from Earth (1.12×10^4 m/s; see Checkpoint 13.22), and so my result is at least not unreasonable. I also note that using my calculated rms speed in Eq. 19.47 yields a value of 1.01×10^5 Pa for P, giving me confidence in my answer.

19.26 A certain number of helium atoms are placed in one container, and an equal number of argon atoms are placed in a separate identical container. The pressure is the same in both containers. Helium atoms have a smaller mass than argon atoms. For which gas, if either, (*a*) is the thermal energy smaller, (*b*) is the average kinetic energy of an atom smaller, (*c*) is the rms speed of an atom smaller?

19.8 Entropy of a monatomic ideal gas

Now that we have derived an expression that gives the thermal energy of a gas in terms of its temperature (Eq. 19.50), we can express the entropy in terms of temperature. Substituting Eq. 19.50 into Eq. 19.32 gives

$$S = \tfrac{3}{2}N \ln(\tfrac{3}{2}Nk_B T) - N \ln\frac{\delta\mathcal{V}}{b}$$

$$= \tfrac{3}{2}N \ln T + \tfrac{3}{2}N \ln(\tfrac{3}{2}Nk_B) - N \ln\frac{\delta\mathcal{V}}{b} \qquad (19.54)$$

or, taking all constant terms together,

$$S = \tfrac{3}{2}N \ln T + \text{constant} \quad \text{(monatomic ideal gas in thermal equilibrium).} \qquad (19.55)$$

This expression describes a gas for which the energy is equipartitioned over the N atoms—that is, in thermal equilibrium. As before, the constant drops out when we consider changes in entropy (Eq. 19.8). For example, if the temperature of the gas is raised from T_i to T_f while the volume is kept constant, the change in entropy is

$$\Delta S = \tfrac{3}{2}N \ln\left(\frac{T_f}{T_i}\right) \quad \text{(constant volume).} \qquad (19.56)$$

EXAMPLE 19.12 Entropy change for a cooling gas

A sample of a monatomic ideal gas consisting of 1.00×10^{23} atoms in a volume of 4.00×10^{-3} m^3 is initially at a temperature of 290 K. The gas is cooled until it reaches a pressure of 5.00×10^4 N/m^2. (*a*) What is the change in entropy for the gas and for each atom? (*b*) What is the ratio of the final rms speed of each atom to its initial rms speed? (*c*) How does the change in the entropy of the gas depend on this ratio $v_{rms,f}/v_{rms,i}$?

❶ **GETTING STARTED** Both rms speed and entropy depend on temperature. As the gas cools, its temperature decreases, and so the rms speed of the atoms and the entropy of the gas both change.

❷ **DEVISE PLAN** I can use the ideal gas law, Eq. 19.51, to determine the final temperature of the gas. Once I know the final temperature, I can calculate the entropy change from Eq. 19.56 and the ratio of final to initial rms speeds from Eq. 19.53. To determine how the change in entropy depends on the ratio of rms speeds, I combine Eqs. 19.53 and 19.56.

❸ **EXECUTE PLAN** (*a*) From Eq. 19.51, the final temperature of the gas is

$$T_f = \frac{P_f V}{Nk_B} = \frac{(5.00 \times 10^4 \text{ N/m}^2)(4.00 \times 10^{-3} \text{ m}^3)}{(1.00 \times 10^{23})(1.38 \times 10^{-23} \text{ J/K})} = 145 \text{ K}.$$

Equation 19.56 then yields for the change in the entropy of the gas:

$$\Delta S = S_f - S_i = \tfrac{3}{2}N \ln\left(\frac{T_f}{T_i}\right) = \tfrac{3}{2}(1.00 \times 10^{23}) \ln\left(\frac{145 \text{ K}}{290 \text{ K}}\right)$$

$$= -1.04 \times 10^{23}. \checkmark$$

Because there are 1.00×10^{23} atoms, the change in entropy per atom is -1.04. ✔

(*b*) From Eq. 19.53 I see that the ratio of final to initial rms speeds is $v_{rms,f}/v_{rms,i} = (T_f/T_i)^{1/2} = [(145 \text{ K})/(290 \text{ K})]^{1/2} = 0.707$. ✔

(*c*) Because $a \ln N = \ln N^a$, I can write Eq. 19.56 as $\Delta S = N \ln(T_f/T_i)^{3/2}$. From part *b* I know that $v_{rms,f}/v_{rms,i} = (T_f/T_i)^{1/2}$, so I can write $T_f/T_i = v^2_{rms,f}/v^2_{rms,i}$. Combining these two expressions, I see that the relationship between the entropy change of the gas and the rms speed ratio of the atoms is $\Delta S = N \ln(v_{rms,f}/v_{rms,i})^3$. ✔

❹ **EVALUATE RESULT** As the gas cools, the rms speed of the atoms decreases, as my answer to part *b* shows. As the rms speed decreases, the number of basic states possible decreases, and so the entropy should decrease, in agreement with my answer to part *a*. Because the system is not closed, there is no reason its entropy cannot decrease, as I have found here.

Equation 19.50, $E_{th} = \tfrac{3}{2}Nk_B T$, tells me that the term on the right in Eq. 19.52 is equal to E_{th}/N, and from this it follows that E_{th} is proportional to v^2_{rms}. Equation 19.32 shows that the entropy is proportional to $\tfrac{3}{2}\ln E_{th} = \ln E_{th}^{3/2}$ and so to $\ln(v^2_{rms})^{3/2} = \ln v^3_{rms}$, in agreement with my result in part *c*.

QUANTITATIVE TOOLS

19.27 (*a*) Does the negative change in entropy in Example 19.12 violate the Eq. 19.5 part of the entropy law? (*b*) What is the change in the thermal energy of the gas in Example 19.12?

We can combine Eq. 19.55 with Eq. 19.7 to obtain an expression that tells us how entropy depends on both volume and temperature. To do so, remember that if we consider both the basic states associated with the spatial distribution of the atoms in a gas and those associated with the velocity distribution, the number of basic states associated with both types of distributions is the product of the number of each type of basic state (Eq. 19.9):

$$\Omega = \Omega_{\text{space}}\,\Omega_{\text{vel}}. \tag{19.57}$$

Taking the natural logarithm of this equation, we have, from $S \equiv \ln \Omega$,

$$S = S_{\text{space}} + S_{\text{vel}}. \tag{19.58}$$

To obtain an expression for the entropy of a gas that provides the dependence on both the volume and the temperature of the gas, we must therefore substitute Eqs. 19.7 and 19.55 into Eq. 19.58:

$$S = N \ln V + \tfrac{3}{2}N \ln T + \text{constant} \quad \text{(equilibrium)}, \tag{19.59}$$

where the constant term $-N \ln \delta V$ in Eqs. 19.7 has been combined with constant terms in Eq. 19.55 to give a new constant term. Equation 19.59 can also be written as

$$S = N \ln(T^{3/2}V) + \text{constant} \quad \text{(equilibrium)}. \tag{19.60}$$

Like Eqs. 19.7 and 19.55, Eq. 19.60 describes a system at equilibrium: If different compartments of the system are not at equilibrium with each other, the entropy of the system is smaller than that given by Eq. 19.60.

Equation 19.60 is in agreement with what we saw in the conceptual part of this chapter: As we increase either the temperature or the volume, the number of basic states—and therefore the entropy—increases. If the system changes from an equilibrium state with temperature T_i and volume V_i to a new equilibrium state with temperature T_f and volume V_f, the change in entropy is

$$\Delta S = \tfrac{3}{2}N \ln\left(\frac{T_f}{T_i}\right) + N \ln\left(\frac{V_f}{V_i}\right). \tag{19.61}$$

Most important, the change in entropy is independent of *how* the system is changed: ΔS depends on only the initial and final equilibrium states of the system. Note that for a system that is not closed, there are no constraints on ΔS: It can be positive, negative, or zero.

19.28 Does the entropy of a gas whose temperature is halved and volume doubled increase, decrease, or stay the same?

Chapter Summary

Entropy of a closed system (Sections 19.1, 19.2, 19.3, 19.5)

Concepts The **probability** of an event is the fraction of times it occurs in a large number of repetitions.

The **macrostate** of a system is that state described by its large-scale properties (on a scale much larger than molecular size), such as volume, pressure, and temperature. The **basic state** of a system is that state described in terms of the properties of the particles that make up the system, such as their position and velocity.

If the interaction between different parts of a system randomizes the distribution of energy in the system, then each particle in the system tends to have an equal share of that energy (**equipartition of energy**), and the particles tend to distribute themselves uniformly in the available volume (**equipartition of space**).

Quantitative Tools If a closed volume of space containing N distinguishable particles is divided into M distinguishable equal-sized compartments, the number of basic states Ω in the system is

$$\Omega = M^N. \tag{19.1}$$

The **entropy** S of a system is a unitless quantity that is a measure of the number of basic states in the system and is defined as

$$S \equiv \ln \Omega. \tag{19.4}$$

For a gas that contains N particles confined to a closed system, if the volume of the gas changes from V_i to V_f, the entropy change of the gas is

$$\Delta S = N \ln\left(\frac{V_f}{V_i}\right). \tag{19.8}$$

The second law of thermodynamics (Sections 19.4, 19.5)

Concepts The **equilibrium state** of a system is the macrostate that has the greatest number of basic states and is therefore the most probable macrostate. At equilibrium, the number of basic states reaches its maximum value Ω_{max}, and the entropy of the system is a maximum.

Second law of thermodynamics (entropy law): A closed system always evolves so as to maximize the number of basic states Ω. When this number has reached a maximum, the system is in equilibrium.

For a system that is not closed, the entropy can increase, decrease, or stay the same.

An **ideal gas** consists of a very large number of particles that occupy a very small fraction of the volume of their container and interact with one another only during collisions that randomize the energy distribution and the spatial distribution of the particles.

Quantitative Tools For a closed system evolving toward equilibrium, the entropy increases:

$$\Delta S > 0. \tag{19.5}$$

For a closed system at equilibrium, $\Omega = \Omega_{max}$ and the entropy does not change:

$$\Delta S = 0. \tag{19.6}$$

These two equations, collectively called the **entropy law,** are a mathematical expression of the second law of thermodynamics.

If a system consists of two subsystems, A and B, the number of basic states for the system is

$$\Omega = \Omega_A \Omega_B \tag{19.9}$$

and the entropy of the system is

$$S = S_A + S_B. \tag{19.10}$$

Entropy and energy (Section 19.6)

Concepts In a monatomic ideal gas, the internal structure of the atoms can be ignored. For this reason, the thermal energy E_{th} of the gas is the incoherent kinetic energy of all the atoms.

Quantitative Tools The **root-mean-square speed** v_{rms} of a gas is the square root of the average of the squares of the speeds of all the atoms of the gas:

$$v_{rms} \equiv \sqrt{(v^2)_{av}}. \tag{19.21}$$

If two gases are in **thermal equilibrium,** their absolute temperatures are the same. In this condition, thermal energy is equipartitioned over the system and the entropy of the system is a maximum.

If a monatomic ideal gas contains N atoms, each of mass m, the average kinetic energy of an atom is

$$K_{\text{av}} \equiv \tfrac{1}{2}m(v^2)_{\text{av}} = \tfrac{1}{2}mv_{\text{rms}}^2 = \frac{E_{\text{th}}}{N}. \tag{19.27}$$

If the thermal energy of a monatomic ideal gas changes from $E_{\text{th,i}}$ to $E_{\text{th,f}}$, the resulting entropy change is

$$\Delta S = \tfrac{3}{2}N\ln\left(\frac{E_{\text{th,f}}}{E_{\text{th,i}}}\right). \tag{19.33}$$

The **absolute temperature** T of a gas, measured in **kelvins,** is related to the rate at which the entropy of the gas changes with respect to its thermal energy:

$$\frac{1}{k_{\text{B}}T} \equiv \frac{dS}{dE_{\text{th}}}, \tag{19.38}$$

where k_{B} is the **Boltzmann constant:**

$$k_{\text{B}} = 1.381 \times 10^{-23}\ \text{J/K}. \tag{19.39}$$

Properties of a monatomic ideal gas (Sections 19.7, 19.8)

Quantitative Tools The pressure in a monatomic ideal gas containing N atoms at an absolute temperature T and occupying a volume V is

$$P = \tfrac{2}{3}\frac{E_{\text{th}}}{V}, \tag{19.48}$$

and the thermal energy of the gas is

$$E_{\text{th}} = \tfrac{3}{2}Nk_{\text{B}}T. \tag{19.50}$$

The **ideal gas law** applied to the gas is

$$P = \frac{N}{V}k_{\text{B}}T. \tag{19.51}$$

The average kinetic energy of its atoms is

$$K_{\text{av}} = \tfrac{1}{2}mv_{\text{rms}}^2 = \tfrac{3}{2}k_{\text{B}}T, \tag{19.52}$$

where m is the mass of each atom, and their root-mean-square speed is

$$v_{\text{rms}} = \sqrt{\frac{3k_{\text{B}}T}{m}}. \tag{19.53}$$

If the gas is in equilibrium, its entropy is

$$S = N\ln(T^{3/2}V) + \text{constant}. \tag{19.60}$$

If the gas changes from an equilibrium state having T_{i} and V_{i} to one having T_{f} and V_{f}, its entropy change is

$$\Delta S = \tfrac{3}{2}N\ln\left(\frac{T_{\text{f}}}{T_{\text{i}}}\right) + N\ln\left(\frac{V_{\text{f}}}{V_{\text{i}}}\right). \tag{19.61}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

Unless told otherwise, assume all gas particles in these problems behave like the particles of a monatomic ideal gas.

19.1 States

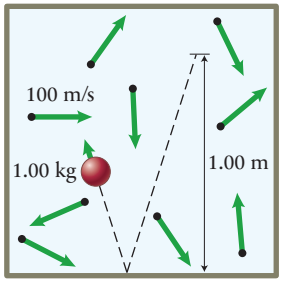
1. With a standard deck of 52 playing cards, what is the probability of randomly picking any (*a*) ace; (*b*) spade; (*c*) black-faced card; and (*d*) an even-numbered card? ••
2. When a pair of standard dice are thrown, what is the probability that the two dice show the (*a*) same number; and (*b*) a pair of odd numbers? ••
3. Particles in a closed container move around and collide with one another and with the walls of the container. There are 5.00×10^{13} basic states possible for this system, and the recurrence time is 365 days. How long is the average time interval between collisions? •
4. A bag contains 13 blue balls, 7 red balls, 10 green balls, and 5 yellow balls. What is the probability of drawing two green balls consecutively? ••
5. A manufacturer sells three types of motorcycles—call them types A, B, and C—each available in yellow, red, glossy black, or matte black and each offered with either a 600-cc motor or an 850-cc motor. (*a*) How many configurations are available? (*b*) If each motorcycle is available with 26 different options packages, how many configurations are available? ••
6. When three standard dice are rolled, find the probability that the sum of the dots on the three top faces will be 5? ••
7. A pendulum is in a container that has been evacuated so that there are only two atoms A and B in the container with the pendulum. If the pendulum initially has six units of energy and the two atoms have none, what is the probability that all six energy units will leave the pendulum and be absorbed by the atoms? ••
8. A pendulum is released from an elevation of 1.15 m and allowed to swing in a room filled with 1.50×10^{21} nitrogen molecules having an average speed of 550 m/s. If the energy of the pendulum is equal to the energy distributed over all the nitrogen molecules, what is the pendulum's mass? The mass of a nitrogen molecule (N_2) is 4.65×10^{-26} kg. ••
9. (*a*) When a coin is tossed five times, which sequence of outcomes is more likely: HHHHH or HTHTH? (*b*) Without regard to order, which is more likely: that in five tosses the coin lands heads-up five times or lands heads-up three times? ••
10. Some musical instruments have valves that can be either opened or closed to produce different tones. How many tones can be produced by a horn that has (*a*) two valves, (*b*) three valves, and (*c*) six valves? (*d*) Some trumpet players do something called *half-valve work*, where they play a valve that is neither fully open nor fully closed. If this technique adds a third possible valve setting, how many tones can be produced with a trumpet that has three valves by a player doing half-valve work? ••
11. Six particles in a container can move in only one of six directions: up, down, left, right, forward, and backward. If the particles move randomly such that each collides and changes direction every 1.0×10^{-6} s, for how many seconds during a 24-h day are all six particles moving in the same direction? •••
12. An adsorbing filter allows gas particles to stick to locations on the filter surface. Once a particle sticks to a location, that location is filled. The filter can no longer remove gas particles when all locations are filled. Each 1.0 nm^2 of the filter surface has six adsorbing locations, each capable of adsorbing one gas particle. In the volume just adjacent to one 1.0-nm^2 area, there are 100 gas particles. Each of these particles has a slightly different energy, making each particle unique. In how many ways can these 100 particles adsorb onto the six locations? •••

19.2 Equipartition of energy

13. A system contains three particles—A, B, and C—and three units of thermal energy. Determine the probability of any one particle having all the energy. •
14. In a system of three particles—A, B, and C—that share nine units of energy, which energy distribution is more likely at a given instant: particle A has two units of energy or particle A has three units of energy? •
15. The air in a room held at a steady temperature T is made up largely of nitrogen molecules ($m_{N_2} = 4.652 \times 10^{-26}$ kg) and oxygen molecules ($m_{O_2} = 5.3135 \times 10^{-26}$ kg). At equilibrium, what is the ratio of the average speed of the nitrogen molecules to the average speed of the oxygen molecules? Even though both gases are diatomic, assume they behave as monatomic gases. ••
16. A closed container holds 35 monatomic A particles that have a combined energy of 90 units. 85 monatomic B particles with an unknown amount of combined energy are then added to the container. If, at thermal equilibrium, each particle have 2 units of energy, how much combined energy do the B particles have? •
17. A certain container holds 50 monatomic particles with a combined energy of 200 units. 25 diatomic particles, with each particle having 2 units of energy, str then added and the system is allowed to reach thermal equilibrium. How many units of energy does each particle in the container have? ••
18. The mass of an oxygen atom is 2.6567×10^{-26} kg. Calculate the rms speed of an oxygen molecule with an average kinetic energy of 1.38×10^{-25} J. ••
19. What is the average kinetic energy of a system made up of 1.00 mol of krypton atoms having an average speed of 125 m/s? The mass of a krypton atom is 1.39×10^{-25} kg. •
20. To a vessel that contains 1.00 mol of monatomic ideal gas A is added 0.100 mol of monatomic ideal gas B. The mass of each atom of gas A is 3.35×10^{-26} kg, and the mass of each atom of gas B is 1.39×10^{-25} kg. What is the final average speed of the A atoms if the final average speed of the B atoms is 123 m/s? ••

21. Inside a box, a 1.00-kg ball is gently dropped from a height of 1.00 m (Figure P19.21). The box contains 1.00 mol of argon atoms, each having a mass of 6.63×10^{-26} kg and moving at an average speed of 100 m/s before the ball is dropped. The ball rebounds elastically against the floor of the box, and so its energy is transferred only to the argon atoms. (*a*) What is the average energy of each argon atom before the ball is dropped? (*b*) After a long time interval has passed, the ball stops bouncing and rests on the box floor. What is the average speed of the argon atoms now? •••

Figure P19.21



22. A spherical balloon that has a diameter of 0.500 m and contains 1.00 mol of helium gas is left in the Sun for 10.0 min. During that interval, the Sun irradiates the balloon with an average intensity of 1120 W/m^2, all of which is absorbed. Before being warmed by the Sun, the average speed of the helium atoms is 367 m/s. What is their average speed after being warmed? Assume the balloon diameter remains constant. The mass of a helium atom is 6.646×10^{-27} kg. •••

19.3 Equipartition of space

23. A cube can be envisioned as an assembly of identical octants (numbered 1–8 in no particular order). The cube contains three identical particles. How many basic states in this system yield a macrostate (*a*) in which all three particles are in octant 1 and (*b*) in which two particles are in octant 8? •

24. A box sitting in a room occupies 1/64 of the room volume. When the box is opened, two identical particles are released into the room. What is the probability that at any instant after they are released, the two particles (*a*) return to the box simultaneously and (*b*) occupy the same 1/64 of the room? •

25. Three particles are released into a box that can be thought of as a set of four identical quadrants. How many basic states are possible if the particles are (*a*) indistinguishable from one another and (*b*) distinguishable from one another? ••

26. Four distinguishable particles move freely in a room divided into octants (there are no actual partitions). Let the basic states be given by specifying the octant in which each particle is located. (*a*) Give an example of a macrostate for this system. (*b*) How many basic states are there? (*c*) The door to this room is opened, allowing the particles to move into an adjacent, identical room, also divided into octants. Now that the amount of space that can be occupied has been doubled, by what factor has the number of basic states increased? ••

27. You have six identical particles in a box divided into four quadrants. For a certain experiment, you need to have a certain number of particles in the upper left quadrant. You observe that in a given day, the distribution of particles you need occurs for 51.43 min. How many particles do you need to be in the upper left quadrant in your experiment? ••

28. A cube that has a volume of 1.00 m^3 contains N distinguishable particles. Determine the probability that all the particles are in a smaller volume V of the cube if (*a*) $N = 1$, $V = 0.500$ m^3; (*b*) $N = 1$, $V = 0.250$ m^3; (*c*) $N = 1$, $V = 0.0100$ m^3; (*d*) $N = 2$, $V = 0.500$ m^3; and (*e*) $N = 4$, $V = 0.500$ m^3. ••

29. A space probe measures the number density of particles of interstellar gas by enclosing 1.00 m^3 of space and measuring the number of atoms that pass from that volume through a detector in the shape of a cylinder 15.0 mm in diameter and 5.00 mm long. In one area of space, the number density of particles of interstellar gas is assumed to be 1 atom/1.00×10^{-6} m^3. Out of a hundred tests, how many times does the detector register the presence of a particle in the cylindrical volume? •••

30. A box divided into four quadrants can hold a maximum of 30 particles in each quadrant. (*a*) If 114 particles are placed in the box, how many basic states and macrostates are possible, if the macrostates are defined as the number of particles in the upper left quadrant? (*b*) What is the most probable number of particles in the upper left quadrant? (*c*) What is the average number of particles in the upper left quadrant? (Hint: Instead of considering where the particles are located, consider where they are not located.) •••

19.4 Evolution toward the most probable macrostate

31. A box divided into compartments A and B contains 14 particles in A and 6 in B. The separating partition allows energy exchange between compartments as the particles collide with the partition. Table 19.2 (page 674) lists the numbers of basic states possible when 10 energy units are distributed in various combinations over the 20 particles. How much more likely is it for compartment A to contain 7 energy units (macrostate $E_A = 7$, $E_B = 3$) than for this compartment to contain 3 energy units (macrostate $E_A = 3$, $E_B = 7$)? •

32. In how many ways can you distribute 10 indistinguishable energy units over 14 distinguishable particles? (Hint: Use Table 19.2, page 674.) •

33. A system has the numbers of macrostates and basic states shown in Table P19.33. (*a*) In which macrostate is the system in equilibrium? (*b*) Which macrostates are least likely? (*c*) How many basic states are possible for the system? (*d*) What is the probability that the system occupies macrostate B? •

Table P19.33

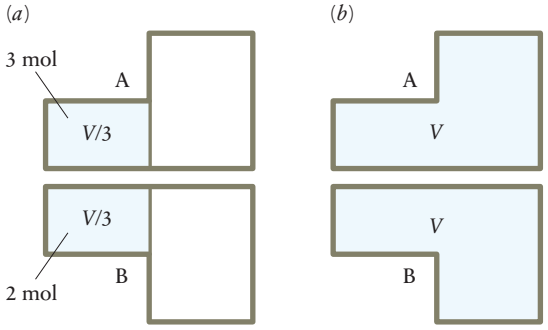
Macrostate	Associated number of basic states
A	1
B	4
C	6
D	4
E	1

34. A box divided into identical compartments A and B contains 25 distinguishable particles in A and 20 distinguishable particles in B. There are nine energy units in the box, initially distributed only among the particles in B. The system is closed, but the energy units can move from one compartment to the other as the particles collide with the separating partition. If one energy unit is exchanged during each partition-particle collision, what is the smallest number of collisions between the particles in B and the partition that must occur before the system reaches equilibrium? ••
35. A box divided into identical compartments A and B contains two distinguishable particles in A and three distinguishable particles in B. There are five energy units in the box, initially shared by the two particles in A. The system is closed, but particles can exchange energy through collisions with the partition. Calculate the number of basic states (*a*) initially and (*b*) once the system has reached equilibrium. ••
36. A box like the one in Problem 31 has two particles in compartment A, two particles in compartment B, and four energy units to be distributed. (*a*) Construct a table like Table 19.2 (page 674) for this setup. (*b*) Does equipartition of energy take place even with such a small number of particles? ••
37. System 1 is in equilibrium, and a separate system 2 is also in equilibrium. The systems are independent of each other but have an equal probability of being in equilibrium. The number of basic states possible in system 1 is 3422 times greater than the number possible in system 2. In system 2, the equilibrium state has 489 possible basic states. How many possible basic states does the equilibrium state of system 1 have? ••
38. A system has three macrostates. Macrostates 1 and 3 are least likely and have one basic state each. Macrostate 2 is the equilibrium state and is six times more likely to occur than either of the other two macrostates. (*a*) What is the probability that the equilibrium state 2 occurs? (*b*) What is the probability that macrostate 1 or 3 occurs? (*c*) What is the sum of the probabilities for all macrostates? ••
39. A cylindrical box is divided into three equal-volume, pie-shaped regions—A, B, and C—by vertical partitions that allow the transfer of energy but not of particles from one region to another. Each region shares a partition with the other two. Each region contains 13 particles, and the system contains 39 energy units, initially all located in region A. If each particle-partition collision transfers one energy unit, is it possible for the system to evolve to its most probable state in (*a*) 18 collisions, (*b*) 27 collisions, (*c*) 29 collisions? •••

19.5 Dependence of entropy on volume

40. What is the entropy of a closed system in which 25 distinguishable grains of sand are distributed among 1000 distinguishable equal-sized compartments? •
41. A number of distinguishable spherical beads, each having a radius of 15.0 mm, are placed into individual compartments in a 1.00-m^3 container. The same number of beads, each having a radius of 20.0 mm, are placed into different compartments in a container of the same volume. Which system has the greater entropy, and by what factor? Assume that each compartment is a cube, with each side equal to the diameter of the bead. •
42. You have 80 particles moving randomly in a closed container, and the system contains 80 energy units. Which is less likely: that a single particle has all 80 energy units, or that each particle has one energy unit? •
43. The volume of a cylindrical chamber is controlled by a movable piston. A sample of gas is placed in the chamber, and the volume is allowed to change from 0.0100 m^3 to 0.100 m^3 at a constant temperature. If the entropy of the system increases by 6.91×10^{18} during the volume change, how many gas particles are there in the chamber? ••
44. A closed box contains 20 particles, 10 on each side of a movable partition. The particles in the left compartment are small and occupy a volume δV_{left}. Those in the right compartment are larger and occupy a volume $\delta V_{\text{right}} = 8\delta V_{\text{left}}$. (*a*) If the box dimension along which the partition moves is 1.00 m long and the partition starts out pushed far to the right, what is the position of the partition when the system reaches equilibrium? (*b*) Does it matter what volume the particles occupy? ••
45. Rank the following systems according to their entropy, lowest first:
    A. 1,000,000 different colors that can be assigned to each of the 70×10^{21} stars in the observable universe
    B. The particles in 1 mol of an ideal gas that can be adsorbed (one particle per site) at any of 83 sites on an adsorbent material
    C. The particles in 0.8 mol of an ideal gas that can be adsorbed (one per site) at any of 106 sites on an adsorbent material
    D. The 100 trillion cells in the human body when each is to be classified as having one of three levels of relative cellular health
    E. The 1×10^{14} microbes in the average human body when each is to be classified as either alive or dead ••
46. Figure P19.46 shows a device consisting of two units, A and B. Each unit has a volume V, and at one end a partition separates off a small compartment that has volume $V/3$, as shown in part *a*. Initially, unit A has 3 mol of a gas in its smaller compartment, and unit B has 2 mol of the same gas in its smaller compartment. When the partitions are removed and the gas in each unit is allowed to expand into the volume V as shown in part *b*, what is the change in the entropy of the two-unit system? ••

Figure P19.46



47. Suppose each of 50 distinguishable objects has a volume of 1.00×10^{-9} m^3 and are initially in a box that has a volume of 1.00×10^{-6} m^3. If the volume of the box is doubled, (*a*) by what factor does the number of basic states change and (*b*) by what factor does the entropy increase? ••
48. A box of volume V has a movable partition separating it into two compartments. The left compartment contains 3000 particles, the right one contains 1000 particles, and initially the partition is in the middle so that the compartments are of equal volume. The partition is released and slides horizontally until the system is in a new equilibrium state. (*a*) In terms of V, what is the volume of each compartment once the new equilibrium state is reached? (*b*) What is the change in the system's entropy during this process? ••
49. A change in the volume occupied by a set of distinguishable particles increases the number of basic states by a factor of 6633 and increases the entropy per particle by a factor of 1.10. How many distinguishable particles are there in the system? •••
50. A box that is 1.00 m long measured from left to right has a sliding partition that separates it into two compartments. The left compartment contains 10 particles, the right one contains 60 particles, and initially the partition is 0.750 m from the left wall of the box. If the partition is allowed to move freely, what is the change in entropy after the system reaches equilibrium? •••

19.6 Dependence of entropy on energy

51. A container holds 1.0 mol of an ideal monatomic gas at 72 K, and then 0.50 mol of another monatomic ideal gas at 126 K is added. What is the equilibrium temperature? •
52. What is the change in entropy of a half a mole of ideal monatomic gas having an initial thermal energy of 75.0 J when this energy is doubled? ••
53. A theory predicts a reaction that produces ideal monatomic gas particles that have the following speeds: 6.2 m/s, 7.4 m/s, 7.4 m/s, 7.8 m/s, 8.3 m/s, 12.6 m/s, 20.1 m/s, 20.1 m/s. What are (*a*) the average speed of the particles and (*b*) their root-mean-square speed? •
54. Three monatomic ideal gas particles have the following velocities (written in component form) in a coordinate system: (3.0, 4.0) m/s, (−6.0, 8.0) m/s, (−5.0, −7.0) m/s. Calculate the (*a*) average speed and (*b*) root-mean-square speed of the particle? ••
55. An ideal monatomic gas is at 125 °C. If you increase the thermal energy of the gas and the root-mean-square speed of the gas particles triples, by what factor does the absolute temperature change? ••
56. Find the rate of change of entropy with respect to thermal energy of a system with an absolute temperature of 100 °C. •
57. The entropy for a certain monatomic gas has been empirically determined to be given by the expression $S = \frac{7}{2}N \ln E_{\text{th}} + NE_{\text{th}}^{2/15}$. Express the factor $1/k_{\text{B}}T$ for this gas in terms of N and E_{th}. ••
58. In (x, y, z) notation, four ideal monatomic gas particles have the velocities $\vec{v}_1 = (-6.0, 5.0, 1.0)$ m/s, $\vec{v}_2 = (4.0, 5.0, -2.0\text{ m/s})$, $\vec{v}_3 = (7.0, 0, 8.0)$ m/s, and $\vec{v}_4 = (-4.0, -9.0, -6.0)$ m/s. (*a*) What is the average velocity of the particles? (*b*) What is the root-mean-square speed of the particles? (*c*) Why is Eq. 19.25, $v_{\text{rms}}^2 = 3(v_x^2)_{\text{av}}$, unlikely to give the correct answer here? ••
59. A box is divided into four equal-sized quadrants. Each quadrant contains gaseous hexane (C_6H_{14}), which we treat as monatomic. The partition that separates the quadrants allows energy to be exchanged, but not particles. However, there is a tiny hole in the partition between quadrants 1 and 2 that allows gas particles to be exchanged. The initial values for the mass of gas in each quadrant and the root-mean-square speed of the particles are Q_1: 3.00 g, 400 m/s; Q_2: 5.50 g, 500 m/s; Q_3: 2.00 g, 420 m/s; Q_4: 6.75 g, 445 m/s. After the system sits for a long time interval, what is the root-mean-square speed for the particles in quadrant 2? •••

19.7 Properties of a monatomic ideal gas

60. Find the rms speed of an oxygen molecule at 0 °C. The mass of an oxygen atom is 2.657×10^{-26} kg? •
61. The molar mass of oxygen (32 g) is slightly greater than that of nitrogen (14 g). But in a tank, the molecules of oxygen have the same rms speed as the molecules of nitrogen in another identical tank. How could this be possible? ••
62. A 1.00-m long cylindrical container with a diameter of 1.20 m contains monatomic ideal gas at 2.00 atm. (*a*) How much thermal energy do the particles of this gas contain? (*b*) If this thermal energy were converted to the gravitational potential energy of a 1400-kg object, how high from the ground must the object be? ••
63. An empty 2.00-L soft drink bottle is filled with 0.0100 mol of a monatomic ideal gas at atmospheric pressure (1.01×10^5 N/m^2). (*a*) What is the average kinetic energy of each gas particle in the bottle? (*b*) Compare the value you obtained in part *a* with the kinetic energy of a bacterium of mass 1.0×10^{-14} kg moving at 1.0×10^{-7} m/s and with the kinetic energy of a 0.0010-kg garden slug moving at 0.0100 m/s. ••
64. A particle accelerator exerts a force on particles that are initially at rest. Each of the particles has a mass of 1.6726×10^{-27} kg, and the point of application of the force is along the entire 100-m length of the accelerator. The particles traverse this distance in 0.0179 s and then collide with a target that is in thermal contact with a monatomic ideal gas. If the particles distribute their kinetic energy elastically, what is the temperature of the gas after the accelerator has been running for a long time interval? ••
65. Very dilute, monatomic cool gases A and B are separated by a partition but allowed to exchange energy through the partition. After a long time interval, $v_{\text{rms,A}} = 6v_{\text{rms,B}}$. What is the relationship between the mass of a particle of gas A and the mass of a particle of gas B? ••
66. A thermally insulated 3.00-L flask contains monatomic xenon gas at an initial temperature of 50.0 °C and pressure of 2.50 atm. You now add enough energy to double the thermal energy of the gas. The mass of a xenon atom is 2.18×10^{-25} kg. (*a*) How many atoms does the flask contain? (*b*) How much energy did you add to the flask? (*c*) By what factor did you change the root-mean-square speed of the atoms? ••

67. In a sample of a monatomic ideal gas that has a number density of 3.61×10^{24} particles/m^3, the average kinetic energy of the gas particles is 7.50×10^{-21} J. If the particles are colliding with a wall that has an area of 1.00×10^{-2} m^2, what is the average change in momentum per unit of time for the gas? ••

68. Chamber X contains monatomic ideal gas X. The gas particles have a root-mean-square speed of 42 m/s. Chamber Y, identical to chamber X, contains a different monatomic ideal gas with the same mass as the gas in chamber X but with twice the number of gas particles as the gas in chamber X. If the absolute temperatures in the chambers are T_X and $T_Y = 1.2T_X$, what is the root-mean-square speed of the particles in Y? ••

69. A jar contains a mixture of helium gas and krypton gas, with the sample containing equal masses of the two gases, $m_{\text{sample,He}} = m_{\text{sample,Kr}}$. Both gases are monatomic, they behave ideally, and the system is in equilibrium. (*a*) What percentage of the pressure in the mixture is due to the helium? What percentage is due to the krypton? (*b*) Explain why your answer in part *a* is reasonable. The atom masses are $m_{\text{He}} = 6.646 \times 10^{-27}$ kg and $m_{\text{Kr}} = 1.391 \times 10^{-25}$ kg. ••

70. Neon atoms at 260 K pass through a fan that gives each mole of neon gas an additional kinetic energy of 16.0 J. (*a*) What is the average temperature of the neon atoms immediately after coming through the fan? (*b*) If you found an increase in temperature in part *a*, how do you explain the fact that fans are usually used to make people cooler rather than hotter? ••

71. A pollution-control device installed in the smokestack of equipment used to run an industrial process measures the root-mean-square speed of pollutant particles given off by the process. If v_{rms} exceeds a given limit, the process is shut down to prevent the release of pollutants into the air. (Consider the particles to be ideal and monatomic.) The safety operator knows that the average mass density of the particles involved in this process is 5600 kg/m^3. (*a*) At what threshold should he set the device if the process is running at 936 °C and the particle diameter is 10 nm? (*b*) As the particles rise and exit the smokestack, the 10-nm particles coalesce to an average diameter of 1.0 μm and cool to 190 °C. What is the root-mean-square speed of the particles as they exit the smokestack? (Ignore viscosity and drag.) •••

72. You are constructing an improvised detector for radon gas, using the cooling core from an old refrigerator capable of cooling air to 255 K, a detector tube 50 mm long, a vacuum pump to evacuate the tube, and an electric circuit that can measure the time interval needed for gas particles to travel a given distance. To determine if there is radon in the air in your room, you intend to inject samples of the air into the evacuated tube and measure the time interval needed for particles in the sample to travel from one end of the tube to the other and their root-mean-square speed for the trip. (*a*) To what scale of measurement should you set the time-interval sensor if you want to detect the monatomic radon atoms? (Ignore any collisions between these atoms and other particles in the sample.) (*b*) Can you suggest any modifications to this device to improve accuracy? (*c*) Do you think you will need to calibrate your system if you have to make the more realistic assumption that the collisions between radon atoms and other particles in the sample cannot be ignored? If so, why? •••

19.8 Entropy of a monatomic ideal gas

73. What is the change in entropy per particle of a monatomic ideal gas heated from 300 K to 400 K? Assume constant volume, and thermal equilibrium in both the initial and final states. •

74. If 1.85×10^5 atoms of a monatomic ideal gas are heated from 300 K to 500 K in a sealed container that has a fixed volume of 0.0560 m^3, what is the change in entropy of the gas? •

75. A monatomic ideal gas in a container of fixed volume contains 4.20×10^{24} atoms. If you add enough thermal energy to double the root-mean-square speed of the gas atoms, by how much have you changed the entropy of the gas? •

76. A container of fixed volume contains 8.72×10^{23} helium atoms, each of mass 6.646×10^{-27} kg. You put the container in a freezer and decrease its temperature from 25.0 °C to −135 °C. What are (*a*) the change in the thermal energy of the gas and (*b*) the change in its entropy? ••

77. A 3.50-L expandable chamber contains monatomic neon gas at 2.00 atm and 30.0 °C. Thermal energy is slowly added to the gas until its volume is 8.00 L while its pressure remains constant. What is the entropy change during this expansion? ••

78. A monatomic ideal gas has its volume halved while its pressure is held constant. What is the change in entropy per atom? ••

79. As a sample of argon gas is heated in a sealed container, its root-mean-square speed changes from 350 m/s to 540 m/s. By how much has the entropy been increased per argon atom? The mass of an argon atom is 6.63×10^{-26} kg. ••

80. A monatomic ideal gas is enclosed in a cubic box of side length ℓ. While preserving the cubic shape, the container volume is decreased until the side length is halved, and this change causes the pressure in the box to triple. (*a*) What is the change in entropy per particle of gas in the box? (*b*) Is this system closed? ••

81. A sample containing 3.65 mol of a monatomic ideal gas is heated from 289 K to 458 K, and the entropy remains constant. If the initial volume of the sample was 0.0980 m^3, by what factor did the pressure increase or decrease during this process? ••

82. A constant-volume encasement vessel surrounds a liquid-nitrogen bath that has cooled 1.00 mol of a monatomic ideal gas to 77.2 K. This system is initially in thermal equilibrium at this temperature, but over time the liquid nitrogen evaporates, and then the ideal gas and the nitrogen gas are in equilibrium in the encasement vessel at room temperature, 297 K. (*a*) What is the change in entropy of the ideal gas as it undergoes this temperature change? (*b*) The nitrogen gas is then condensed and again cools the ideal gas. What is the entropy change for the ideal gas as it returns to 77.2 K? (*c*) Where

in this process did entropy increase, as required by the second law of thermodynamics for an irreversible process? ••

83. Ten particles of a monatomic ideal gas are initially at thermal equilibrium with their surroundings. When the sample is then heated to 634 K while the volume is held constant, the entropy doubles. What was the initial temperature of the sample? ••

84. In a cylindrical chamber that initially has a volume of 2.25 L, a sample of helium gas is at equilibrium at 3.20 atm and 85.0 °C. The gas is slowly compressed to a new equilibrium state at a volume of 1.10 L and pressure of 4.40 atm. What is the entropy change during this compression? ••

85. A thermally insulated chamber is divided by a partition into two compartments of equal volume. The left side contains 9.00×10^{23} helium atoms, and the right side contains 2.70×10^{24} argon atoms. The gases in both compartments are in equilibrium and at the same temperature. A small hole develops in the partition so that after a long time interval both gases leak through the hole and eventually the system comes to a new equilibrium state. (*a*) What is the difference in entropy between the two equilibrium states? Which state, if either, has the greater entropy? (*b*) What is the change in the thermal energy of the gases during this process? Assume ideal behavior for all atoms. ••

86. A rigid 1.20-L bottle contains helium gas at 150 °C and 2.75 atm. If you slowly add 755 J of thermal energy to the gas, what is the entropy change for the gas? Assume the helium gas is ideal. ••

Additional Problems

87. Suppose you are dropping five small pieces of clay randomly onto a chessboard (eight squares by eight squares). If basic states are determined by the squares on which each piece of clay lands, how many basic states are possible? Assume a square can be occupied by multiple pieces. •

88. Consider two egg cartons, one with spaces for a dozen eggs and one with spaces for a half-dozen eggs. If each carton contains six eggs, all of different colors, what is the ratio of the number of possible basic states in the one-dozen carton to the number of possible basic states in the half-dozen carton? •

89. Two boxes each have a volume of 1.50 m^3. Box A contains 1000 helium atoms, and box B contains 2000 neon atoms. When the boxes are connected together by a very small tube, the atoms redistribute themselves throughout both boxes. What is the entropy change of the two-box system during the redistribution? •

90. A sample of neon gas is at 100 °C. The mass of a neon atom is 3.35×10^{-26} kg. (*a*) What is the average kinetic energy of a neon atom? (*b*) If the thermal energy of the sample is 175 J, how many neon atoms does it contain? •

91. At 35 °C, the root-mean-square speed of a certain monatomic ideal gas is 186 m/s. (*a*) What is the mass of each atom? (*b*) What is the average kinetic energy of a typical atom at that temperature? •

92. A 0.00126-kg sample of an ideal gas in a fixed-volume, 0.750-L container at 50.0 °C exerts a pressure of 2.24 atm. What is the mass of one particle of this gas? ••

93. The atoms of an element can come in different forms, depending on the number of neutrons in the nucleus. These different forms are called *isotopes* of the element. The two most common isotopes of uranium are uranium-235 (143 neutrons) and uranium-238 (146 neutrons), having masses of 3.90×10^{-25} kg and 3.95×10^{-25} kg, respectively. In a sample of uranium gas, these isotopes have slightly different root-mean-square speeds, and this difference can be used to separate the isotopes for use in nuclear reactors or nuclear weapons. When both gases are at the same temperature, what is the percent difference in the root-mean-square speeds of the atoms of each isotope? Which isotope has the greater root-mean-square speed? ••

94. A sealed 1.50-L chamber filled with helium gas initially at 20 °C and 1.00 atm is heated until the gas temperature is 232 °C. (*a*) How much thermal energy is added to the gas during this process? (*b*) What is the entropy change of the gas during this process? (*c*) What was the final pressure inside the container? Ignore any thermal expansion of the container; the mass of a helium atom is 6.646×10^{-27} kg. ••

95. (*a*) At what temperature is the root-mean-square speed of neon atoms (mass 3.35×10^{-26} kg) equal to the speed of sound in a typical room, 344 m/s? (*b*) At the temperature you found in part *a*, how much thermal energy do 6.02×10^{23} neon atoms have? ••

96. A monatomic ideal gas is in a fixed-volume flask. (*a*) By what factor must you increase the root-mean-square speed of the gas particles to double the gas pressure? (*b*) By what factor does the thermal energy of the gas change when you increase v_{rms} by this factor? ••

97. A sample of a monatomic ideal gas contains N atoms. You want to compress the gas slowly to one-fourth of its original volume. (*a*) By what factor must you change the absolute temperature of the sample in order to increase its entropy by $3N$? (*b*) For the process in part *a*, by what factor does the root-mean-square speed of the gas particles change? ••

98. The temperature of the Sun's corona (the outermost gas layer) is 1.0×10^6 K, and this layer is just above the photosphere layer, which has a radius of 6.96×10^8 m. The mass of the Sun is 1.99×10^{30} kg. Is the root-mean-square speed of helium atoms in the corona great enough to allow the atoms to escape from the Sun's gravitational pull? The mass of a helium atom is 6.646×10^{-27} kg. ••

99. Consider a helium atom that is part of Earth's atmosphere and is initially at the planet's surface, where the temperature is 20 °C. To what maximum altitude can this atom rise if its speed is equal to the root-mean-square speed of the helium in the atmosphere? The mass of a helium atom is 6.646×10^{-27} kg, the mass of Earth is 5.97×10^{24} kg, and Earth's radius is 6.38×10^6 m. (Careful! The value of g varies with altitude.) ••

100. A box initially divided into two equal halves by a partition contains some number of particles. When the partition is removed, the particles are free to move throughout the volume of the box. Assume that all the

particles have an average speed that allows them to cross from one side of the box to the other every 2.50 s, on average. Note that this does not mean that they uniformly switch sides; it means that they move randomly in such a way that they enter a new random distribution every 2.50 s. Determine the probability of all particles being in the left half of the box and the average time interval required for this to happen for (*a*) two particles and (*b*) ten particles. ●●

101. A cylindrical chamber is oriented with its long central axis along the vertical *y* axis of an *xyz* coordinate system. The top face of the chamber is a piston that can slide up and down to change the chamber volume. The chamber is initially filled with a monatomic ideal gas. When the chamber temperature is lowered, the piston slides downward very slowly at constant velocity until the chamber volume is one-fifth of its original value. At this instant, the contraction stops with the gas in an equilibrium state. (*a*) By what factor does the root-mean-square speed of the gas particles change during this contraction? Does this speed increase or decrease? (*b*) By what factor does the thermal energy of the gas change during this contraction? Does the energy increase or decrease? ●●●

102. A cylindrical canister fitted with a piston at one end contains an ideal gas and floats at the surface of a freshwater lake, where the temperature is 300 K. The gas pressure inside the canister is just enough to balance atmospheric pressure, and so the piston does not slide downward. Initially the canister volume is 3.64 L. The canister is lowered into the lake until it is 12.0 m below the surface, where the temperature is 280 K. (*a*) What is the canister volume once equilibrium is reached at this depth? (*b*) What is the change in entropy of the gas during this process? Ignore friction and ignore the effect of gravity on the piston. ●●●

103. Your work in a laboratory requires a certain dangerous ideal gas, which you will be using for the next three weeks. You know that the gas is corrosive and that gas particles are constantly colliding with one another and, more worrisome, with the walls of the container. Thus there is some finite probability that a gas particle will react with a particle of the container and weaken the container wall. You have 4.00×10^{-3} kg of this gas in a spherical 6.45-L container in which the walls are 10.0 mm thick. The mass of each gas particle is 6.054×10^{-26} kg. The temperature of the gas is always close to room temperature and certainly never rises above 30 °C. After a little online search, you realize that each time a gas particle hits the container wall there is a one in 3 million chance of a corrosive reaction. That causes you to wonder about safety issues. ●●● CR

104. A tank filled with helium gas has been left in a hot car for hours. The mass of the filled tank is 21.2 kg, and after you use all the helium to fill 100 balloons for a child's birthday party, the tank mass is 20.8 kg. You notice that the filled balloons are shrinking over time. Each filled balloon is roughly spherical and started out with a radius of 200 mm, but now each radius has shrunk to 100 mm. Some of the children are afraid the balloons will deflate completely. You estimate that the gauge pressure inside the balloon was initially 50.0 kPa and that this pressure must increase linearly with balloon surface area. You are puzzled by the shrinking because the air temperature at the party is 35.0 °C. That causes you to wonder whether the balloons have an equilibrium radius or you should prepare the children for complete deflation. ●●● CR



20

Energy Transferred Thermally

20.1 Thermal interactions
20.2 Temperature measurement
20.3 Heat capacity
20.4 *PV* diagrams and processes

20.5 Change in energy and work
20.6 Isochoric and isentropic ideal gas processes
20.7 Isobaric and isothermal ideal gas processes
20.8 Entropy change in ideal gas processes
20.9 Entropy change in nonideal gas processes

CONCEPTS

QUANTITATIVE TOOLS

We use terms like *hot* and *cold*, *heating* and *temperature* based on our everyday interpretations of these words. We know from experience that to raise the temperature of water to the boiling point for a cup of tea, we need to heat the water, which requires energy. The branch of physics that deals with these concepts, called *thermodynamics* (from the Greek *therme*, meaning "heat," and *dunamis*, meaning "power"), was developed in the early 19th century to describe the working of steam engines. The concepts of entropy and absolute temperature introduced in Chapter 19 were developed as a consequence of that work. In this chapter we connect the statistical ideas from Chapter 19 to the macroscopic concepts of temperature, pressure, and volume.

20.1 Thermal interactions

The concept of temperature is familiar: It's a measure of "hotness." According to what we saw in Chapter 19, if two objects that are not at the same temperature are placed in contact with each other, energy is transferred from the object at the higher temperature to the object at the lower temperature until the temperatures of the two are equal. We attribute this transfer of energy to a **thermal interaction** between the two objects at different temperatures. As we saw in Chapter 19, this thermal interaction is the overall effect of countless interactions at the atomic scale.

Let us begin by considering the familiar situation shown in **Figure 20.1**: heating a pot of water on a stove. As we heat the water, energy flows from the flame to the pot and then from the pot to the water. The flame, which is hotter (that is, at a higher temperature) than the pot, heats the pot, and the pot, which is hotter than the water, heats the water. If we consider the water as our system and make an energy diagram for that system, we obtain the result shown in **Figure 20.2**. As the water gets hotter, its thermal energy increases. The burning flame provides that energy, but it is not part of the system, and so there is no change in our system's source energy. The system's kinetic energy and potential energy do not change either, and so the system's energy increases because of the increase in thermal energy.

Figure 20.1 When a gas burner heats a pot of water, energy flows from higher temperature to lower temperature.

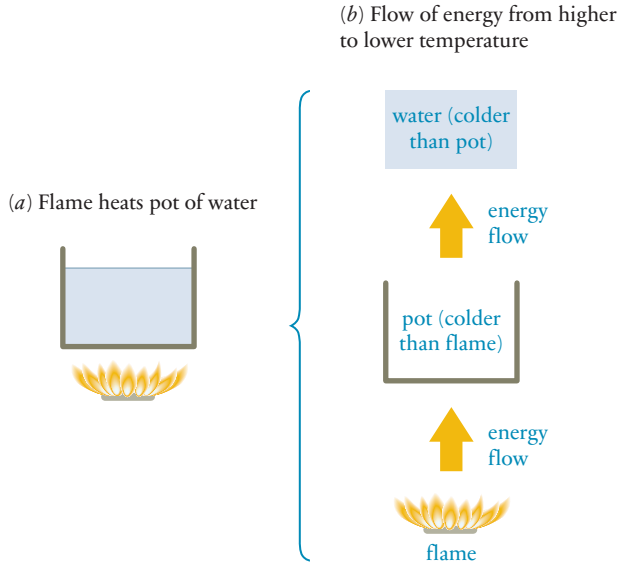
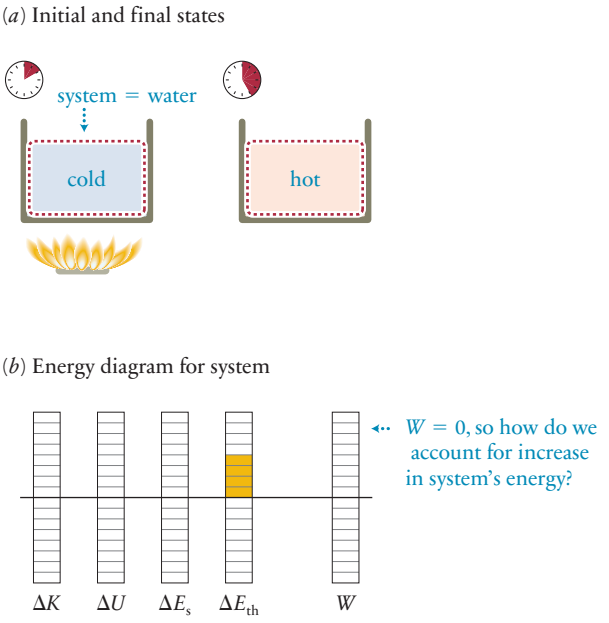


Figure 20.2 When a pot of water is heated, the energy diagram for a system comprising only the water is unbalanced.



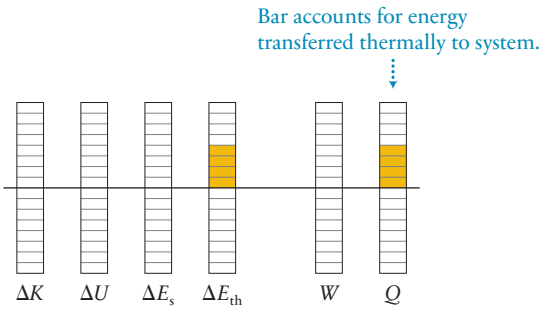
How do we account for this increase in the energy of the system? Although there are external forces exerted on the water—the force of gravity and a normal force exerted by the pot on the water—neither force causes any displacement, and so does no work on the system.

Consequently the energy diagram of Figure 20.2 is unbalanced: The energy changes are not equal to the work done on the system. The reason for this unbalance is that this energy diagram does not account for the energy transferred into the system by the heating. To make the diagram complete, we need to add a bar that represents the energy transferred into the system by the thermal interactions among the flame, pot, and water. The amount of energy transferred by thermal interactions ("thermally") is denoted by the letter Q (**Figure 20.3**). The law of conservation of energy requires that the change in the energy of the system be equal to the energy transferred into it, so the height of the bar representing Q in Figure 20.3 has to be equal to the height of the bar representing the increase in thermal energy.

20.1 (*a*) What are the SI units of Q? (*b*) For the process depicted in Figure 20.2*a*, make an energy diagram for each of these systems: (*i*) water, pot, and flame; (*ii*) pot and flame; (*iii*) pot.

CONCEPTS

Figure 20.3 To complete the energy diagram of Figure 20.2, we must add a bar to account for the energy transferred thermally into the system (represented by the symbol Q).



The quantity Q is frequently called *heat*, a term we shall avoid because it is easy to confuse the verb *to heat* with the noun and to think of "heat" as a property of an object instead of a quantity of energy transferred across a boundary. When it is necessary to refer to Q by name, we shall use the term **energy transferred thermally,** reserving the verb *to heat* to refer to the process of increasing the thermal energy of a system. (For example, *I am heating a pot of water* means that I am increasing its thermal energy.)

Now consider the situation in **Figure 20.4*a***, where two gases are separated by a fixed partition that is not thermally insulating. Initially gas A in the left compartment is at a lower temperature than gas B in the right compartment. Because the partition is not thermally insulating, the gases interact thermally, and energy is transferred thermally from the hotter gas to the colder one until their temperatures are the same. The energy transferred thermally is represented by the arrow labeled Q. The energy diagram for gas A in Figure 20.4*b* consequently shows an increase in thermal energy and the energy transferred thermally, Q, is positive.

20.2 Make an energy diagram for gas B in Figure 20.4.

As Checkpoint 20.2 shows, the sign of Q depends on the direction of the energy transfer. Just as for work, if the transfer is into the system (that is, the energy of the system increases), Q is positive (Figure 20.4*b*); if the transfer is out of the system, Q is negative (Checkpoint 20.2).

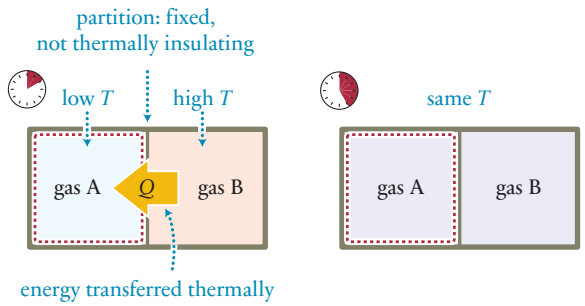
Let us examine in more detail the difference between transferring energy by doing work on a system (that is, through mechanical interactions) and by transferring energy thermally to it. In **Figure 20.5*a*** a spring-loaded piston compresses an ideal gas into a smaller volume, doing work on the gas. We shall ignore any friction between the piston and the container walls. The ideal gas is *thermally insulated;* that is, it is not interacting thermally with its surroundings. A process that does not involve any thermal transfer of energy is called an **adiabatic process** ($Q = 0$).

Because the process depicted in Figure 20.5*a* is adiabatic, the amount of energy transferred to the gas is equal to the work done on the gas by the piston and is determined by the force $\vec{F}^{c}_{pg}$ exerted by the piston on the gas and the force displacement $\Delta\vec{x}_F$.

How is the energy of the gas raised? There are no changes in the kinetic energy of the center of mass of

Figure 20.4 When gases of different temperatures are placed in thermal contact, energy is transferred thermally from the hotter to the cooler gas until the temperatures of the two gases are the same.

(*a*) Gases at different temperatures are placed in thermal contact



(*b*) Energy diagram for gas A

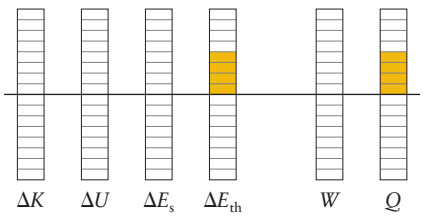
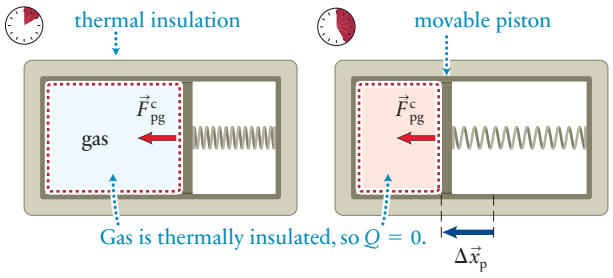
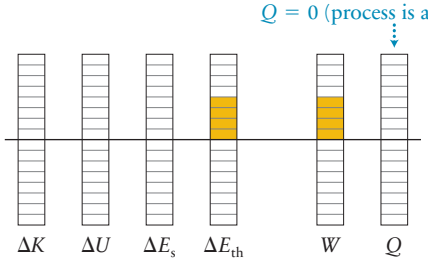


Figure 20.5 A piston does work on a thermally insulated ideal gas, increasing the gas's temperature.

(*a*) Spring-loaded piston compresses thermally insulated ideal gas

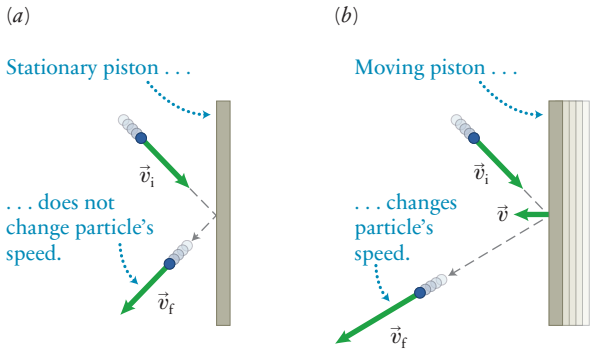


(*b*) Energy diagram for gas



CONCEPTS

Figure 20.6 A stationary piston does not change the speed of a gas particle that collides with it, but a moving piston does.



the gas because the gas as a whole remains at rest. There are also no changes in the potential energy of the gas because there is no potential energy associated with the interactions between particles in an ideal gas. Nor are there any changes in the source energy. Therefore the work W done on the gas must raise the thermal energy of the gas (Figure 20.5*b*). We can understand how this happens by considering the (elastic) collisions of the gas particles with the piston. If a particle collides with the stationary piston (**Figure 20.6*a***), the particle's speed is unaltered by the collision. When the piston is moving toward the incoming particle, however, the particle gains speed because of the collision and therefore its kinetic energy increases, just as a tennis ball bounces faster off your racquet when you swing it toward the ball. So, as the piston moves in, the particles colliding with it speed up (Figure 20.6*b*), increasing their incoherent kinetic energy and thus the thermal energy of the gas.

Note that in Figure 20.5*a* the volume decreases. During this decrease, the force exerted by the piston on the gas and the force displacement are in the same direction. Consequently the work done on the gas is positive. Had the piston moved in the opposite direction (expanding the volume), the work done on the gas would have been negative.

> **When the volume of a gas decreases, the work done on the gas is positive. When the volume increases, the work done on the gas is negative. When there is no change in volume, the work done on the gas is zero.**

Summarizing the information from Figures 20.4 and 20.5, we see that there are two ways to transfer energy to a system: mechanically (by doing work on the system) and thermally (see the box "Transferring energy to a system"). Both types of transfer can occur simultaneously. In that case the change in the energy of the system ΔE is equal to the sum of the work done on the system, W, and the energy transferred thermally to it, Q.

Let us now turn our attention to the transformation of a system from one macrostate to another. Such a transformation is a called a **process,** and if the process is carried out in such a way that the system remains very close to equilibrium at all instants during the process, the process is said to be **quasistatic.**

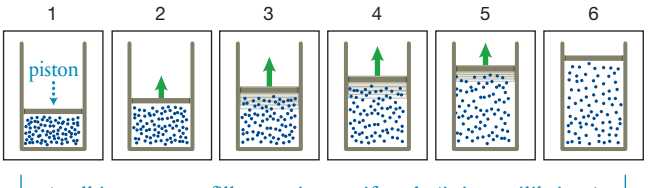
CONCEPTS

Transferring energy to a system

	DOING WORK	TRANSFERRING ENERGY THERMALLY
Technical term:	Mechanical interaction between system and its environment	Thermal interaction between system and its environment
Mediated by:	Macroscopic pushes and pulls	Atomic and/or molecular collisions
Requires:	A nonzero external force $\vec{F}_{ext}$ exerted on the system and a nonzero force displacement $\Delta\vec{x}_F$	A nonzero temperature difference ΔT between the system and its environment
Amount of energy transferred *into* system:	W ("work done on system")	Q ("energy transferred thermally to system")
Sign:	$W > 0$ when external force and force displacement are in the same direction (energy of system increases in absence of other transfers)	$Q > 0$ when environment is at a higher temperature than system (energy of system increases in absence of other transfers)
	$W < 0$ when external force and force displacement are in opposite directions (energy of system decreases in absence of other transfers)	$Q < 0$ when environment is at a lower temperature than system (energy of system decreases in absence of other transfers)
Equilibrium:	The vector sum of the external forces is zero, and the force displacement of each external force is zero (mechanical equilibrium)	System is at same temperature as its environment (thermal equilibrium)

Figure 20.7 Quasistatic versus non-quasistatic expansion of a cylinder containing a gas.

(*a*) Quasistatic expansion: piston rises gradually



(*b*) Non-quasistatic expansion: piston rises suddenly

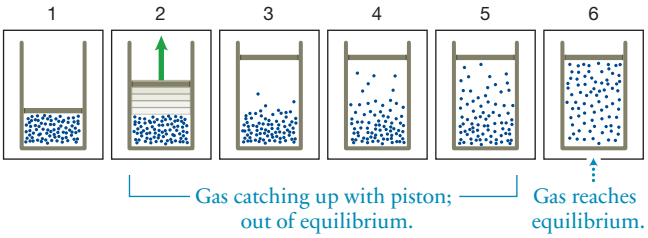


Figure 20.7*a* shows the expansion of a volume of gas in a cylinder. To make the expansion quasistatic, the piston is moved slowly so that the gas always uniformly fills all the available space, and each particle has its equipartition share of the thermal energy. In each frame shown in the drawing, the gas is in a different state, but each state is an equilibrium state. In a quasistatic process, therefore, the pressure may change during the process, but at any given instant the pressure has a uniform, well-defined value at all locations in the volume of gas. The same is true of the temperature: The final temperature may differ from the initial temperature, but at any given instant the temperature has a uniform, well-defined value at all locations in the volume of gas. There are no locations where the pressure is greater and no spots that are hotter than others. In Figure 20.7*a*, the entropy changes as the gas expands (because the entropy of an ideal gas changes with volume and temperature, Eq. 19.60), but at each instant it has the maximum value for that equilibrium state.

Suppose, however, that the piston moves at a rapid speed, as illustrated in Figure 20.7*b*. The motion of the piston is so fast that the gas can't keep up with it. In frame 2, for example, the gas still mostly occupies the original volume. Consequently, the pressure at the bottom of the cylinder is still nearly the same as it was in frame 1, while it is nearly zero at the top, where we still have a vacuum. Although the piston reaches its final position in frame 3, it takes the gas another three frames before it is evenly distributed and in equilibrium. In frames 2–5 of Figure 20.7*b*, the gas is not in equilibrium, and so the entropy is not maximum. Because particles in a gas typically move at very great speeds, they redistribute themselves rapidly. Thus, in practice, even if a piston such as the one in Figure 20.5 moves rapidly, the resulting process can still be quasistatic.



20.3 Suppose you were to play the two film clips shown in Figure 20.7 backward. Would the resulting processes be possible?

Let us now examine how the entropy changes when energy is transferred to a system in a quasistatic process. First consider the process shown in Figure 20.4, where energy is transferred thermally between two gases. The energy transferred thermally to gas A is positive ($Q > 0$). The temperature of gas A increases. Therefore by Eq. 19.56 the entropy of the gas increases. Conversely, Q is negative for gas B, the temperature of B decreases, and so its entropy decreases. Consequently, the sign of the entropy change of a system is the same as the sign of Q for that system:

> ***Transferring energy thermally to a system increases its entropy. Transferring energy thermally from a system decreases its entropy.***

Next, let us determine what happens to the entropy of a system when work is done on it in the absence of any thermal interaction ($Q = 0$). Consider, for example, the quasistatic compression of the thermally insulated gas in Figure 20.5 again. As the piston does work on the gas, the thermal energy of the gas goes up, but at the same time the volume goes down. In Chapter 19 we saw that an increase in the thermal energy of a system raises its entropy. A decrease in the system volume, however, lowers the entropy, so we can't say offhand what happens to the entropy of the gas as it is compressed.

Let us therefore consider the closed system consisting of the contents of the box in Figure 20.5*a*. In contrast to the equilibration of temperatures in Figure 20.4, which is an irreversible process, the compression of the gas and the expansion of the spring are reversible. Because the process is quasistatic, the system remains in equilibrium at all instants during the compression, and therefore the entropy law tells us that the entropy of the (closed) system doesn't change: $\Delta S = 0$ (Eq. 19.6).

Because we are interested in the process (which takes place in the gas), we divide the system into two parts: the gas and the spring-loaded piston ("piston" for short). Because entropy is an extensive quantity, we can write the change in the system's entropy as $\Delta S = \Delta S_{\text{gas}} + \Delta S_{\text{piston}} = 0$.

In the spring-loaded piston, potential energy is converted to kinetic energy, both of which are coherent forms of energy (see Figure 7.10). Regardless of the speed of the piston and the configuration of the spring, there is only one way to distribute the energy and atoms such that all the energy is coherent. At any instant, therefore, the number of basic states of the spring-loaded

piston is 1, and so the entropy of the spring-loaded piston does not change: $\Delta S_{\text{piston}} = 0$.

If $\Delta S_{\text{piston}} = 0$ and $\Delta S = 0$, then the entropy of the gas cannot change either: $\Delta S_{\text{gas}} = 0$. So we can conclude that in a quasistatic process during which we do work on a gas without thermally transferring any energy to it ($Q = 0$), the entropy of the gas does not change. Generalizing this idea to any type of system undergoing a quasistatic adiabatic process, we can state:

The entropy of a system undergoing a quasistatic adiabatic process does not change.

In a quasistatic adiabatic process, the temperature and volume changes are coordinated in such a way as to keep the entropy constant. This conclusion holds for only *quasistatic* adiabatic processes. For instance, in Example 19.9, a gas adiabatically expands into a larger volume, increasing the entropy of the gas. The expansion is adiabatic, but not quasistatic. Therefore the entropy of a system undergoing a non-quasistatic adiabatic process can change.

20.4 Which of the following systems undergo a quasistatic process? (*a*) A gas-containing balloon is popped inside a vacuum container. System: container, gas, and balloon. (*b*) A cup of coffee cools down. System: coffee. (*c*) A sample of compressed gas in a cylinder fitted with a heavy piston is allowed to expand, pushing the piston outward. System: gas. (*d*) An ice cube melts in a glass of hot water. System: water and ice cube.

20.2 Temperature measurement

In Chapter 19 we defined the absolute temperature as a rather abstract concept related to the rate at which entropy changes with changes in thermal energy. Entropy is a measure of the number of available basic states in a system, and thermal energy is the energy associated with the incoherent internal motion of the system. Neither of these quantities can be measured directly. We must therefore figure out some convenient way to measure the temperature of an object.

Most of the methods that have been developed to measure temperature rely on measuring some physical quantity that varies with temperature. For example, in a glass thermometer, illustrated in Figure 20.8, the thermal expansion of a liquid provides a means of measuring temperature. A glass tube is connected to a bulb filled with glycerin, mercury, or some other liquid whose mass density decreases with rising temperature. Because the mass of the liquid is constant, its volume increases and therefore the temperature can be determined by measuring the height of the liquid in the glass tube. To calibrate this type of thermometer, it is placed in melting ice and allowed to come to thermal equilibrium. Once equilibrium is reached, a mark is made on the glass tube at the height of the liquid in the tube. The thermometer is then placed in boiling water and allowed to come to thermal equilibrium. The tube is marked at the new height of the liquid in the tube. The two marks on the tube are assigned some standard values representing the freezing and boiling points of water—historically 0 °C and 100 °C, respectively.

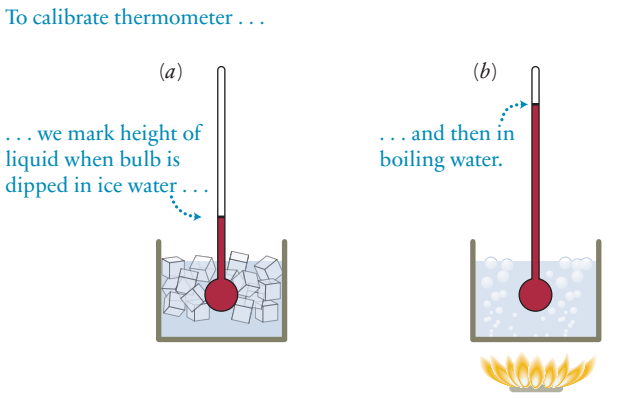


Figure 20.8 Calibrating a gas thermometer by immersing its bulb in ice water and then in boiling water.

If a liquid's change in volume is proportional to its change in temperature, the temperature of any substance can be measured by immersing the thermometer in the substance and noting how far up the glass tube the enclosed liquid rises. If the thermometer is immersed in a cup of milk and the liquid in the tube rises, say, 27/100 of the distance between the 0 °C and 100 °C marks, the temperature of the milk is 27 °C. Of course, in commercial thermometers there is no need to measure heights in fractions because the glass tube is marked all along its length and the temperature is read directly.

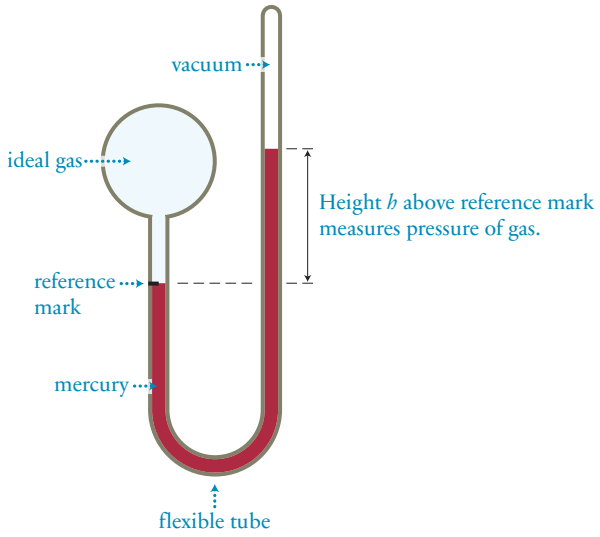
Is the expansion of a liquid proportional to the absolute temperature we defined in Chapter 19? No—as a matter of fact, the expansion of one liquid isn't even proportional to that of another liquid. For this reason, glass thermometers are accurate over only a relatively narrow range of temperatures.

How, then can we measure temperature accurately? In Section 19.7 we derived the ideal gas law, $P = Nk_{\text{B}}T/V$, which relates the absolute temperature to two macroscopic quantities that are easily measured: pressure and volume. This law is a direct consequence of our definition of absolute temperature and is valid for an ideal gas, that is, it is valid as long as the pressure, volume, and absolute temperature are such that we can ignore the internal structure of the gas particles and the interactions between particles (except for collisions between them). No gas is ideal, but many gases can approximate ideal behavior. For a dilute gas of helium atoms, for instance, the conditions for ideal behavior are satisfied down to temperatures that are quite low.

Figure 20.9 illustrates the basic principle of an **ideal gas thermometer**, which measures the pressure in a gas

CONCEPTS

Figure 20.9 An ideal gas thermometer.



sample kept enclosed in a fixed volume. A bulb filled with a gas at low pressure is connected to a thin, flexible U-shaped tube that contains mercury. The end of the tube opposite the gas-filled bulb is capped and evacuated. The difference in height between the levels of mercury in the two arms of the tube determines the pressure in the gas in the bulb (see Section 18.6). For a fixed value of V, the ideal gas law, $P = Nk_BT/V$, tells us that the pressure P is proportional to the absolute temperature T.

Suppose now that the absolute temperature of the gas in the bulb is raised. The pressure in the gas increases, pushing the mercury in the left arm down below the reference mark and increasing the volume V of the gas, as illustrated in **Figure 20.10*a***. To restore the gas to its original volume, we raise the right arm of the flexible U-tube until the mercury level in the left arm is back at the reference mark (Figure 20.10*b*).

Provided the mercury level in the left arm is always kept at the reference mark, the volume V of the gas is fixed and the height h of the mercury in the right arm measures the pressure in the gas at a constant volume. Under these conditions, the height h is a direct measure of the pressure and therefore the absolute temperature T. The ideal gas law tells us that at $T = 0$ the pressure is zero, which means there is a vacuum above both arms of the U-shaped tube. So we know that at $T = 0$ the height h should be zero.

To calibrate the thermometer, we need to find some reproducible thermal phenomenon and assign it a fixed value. Until 2019, the standard of absolute temperature was based on the **triple point of water**—the unique conditions under which liquid water, solid ice, and water vapor (a gas) coexist in thermal equilibrium. To relate absolute temperature to historical definitions of temperature, the absolute temperature at the triple point of water is set at 273.16 K:*

$$T_{tp} \equiv 273.16 \text{ K}.$$

Once the height of the mercury column in an ideal gas thermometer has been measured at the triple point of water, we can draw a straight line between the origin at $T = 0$, where we know h should be zero, and the triple point. We can then determine any other absolute temperature simply by reading off the absolute temperature corresponding to the measured height (**Figure 20.11**). The important point is that the curve is linear: As long as the gas behaves like an ideal gas and is kept at constant volume, the pressure (and therefore the height of the mercury column) is directly proportional to the absolute temperature.

Figure 20.10 Using an ideal gas thermometer to measure temperature.

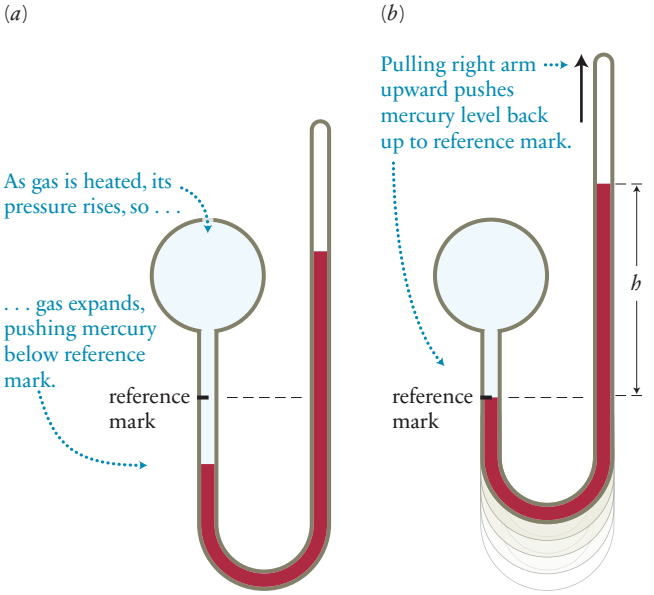
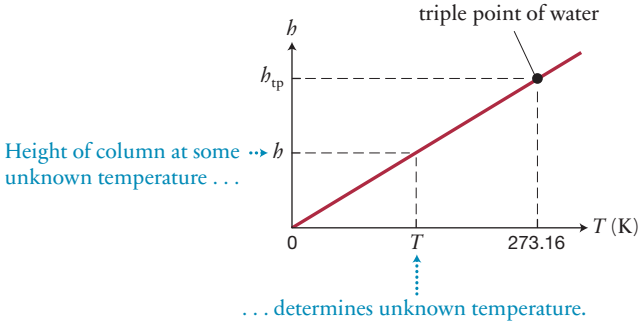


Figure 20.11 The height-versus-temperature graph is linear for an ideal gas thermometer. The thermometer is calibrated at the triple point of water. Any temperature can then be read from the graph.



*The triple point of water occurs at a temperature of 0.01 °C (roughly 32 °F) and a pressure of 610 Pa (about one-hundredth of atmospheric pressure).

CONCEPTS

Temperature measurement scales

Methods for measuring and defining temperature were developed long before the development of the gas thermometer and the concept of absolute temperature. Starting in the early 18th century, several scales were defined for measuring temperature, two of which are still in use. The most widely used scale is the *Celsius scale*. Temperatures on this scale are measured in units that are called *degrees Celsius* (°C) and that have the same magnitude as the kelvin (a temperature change of 1 °C is the same as a temperature change of 1 K). The zero of the Celsius scale is the freezing point of water, and the boiling point of water (at a pressure of 101.325 kPa) is set at 100 °C.

We can convert between the Celsius and Kelvin scales as follows. If T_C represents the temperature expressed in degrees Celsius and T the absolute temperature expressed in kelvins, we have

$$T_C = \frac{1\ °C}{1\ K} T - 273.15\ °C. \qquad (1)$$

The older *Fahrenheit scale* is used in the United States. Temperatures on this scale are measured in units called *degrees Fahrenheit* (°F). On this temperature scale, the freezing point of water is 32 °F and the boiling point is 212 °F, placing the boiling and freezing points of water exactly 180 °F apart. If T_F is the temperature expressed in degrees Fahrenheit, the relationship between the Celsius and the Fahrenheit scales is

$$T_C = \frac{5\ °C}{9\ °F}(T_F - 32\ °F). \qquad (2)$$

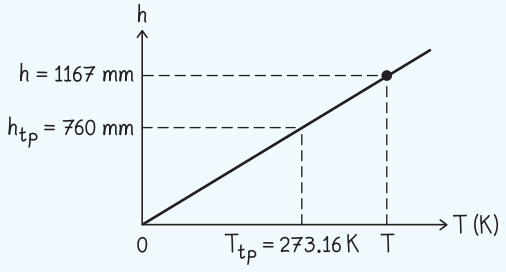
The triple point of water had been used to define the kelvin, but this unit was redefined in terms of the Boltzmann constant (see Section 19.6), which is fixed at exactly 1.380649×10^{-23} J/K. Notice that no degree symbol, °, is used with kelvins: 13.0 °C, 55.4 °F, but 286 K. Readings in degrees Celsius or Fahrenheit can be negative, but readings in kelvins are always positive.

EXAMPLE 20.1 Measuring absolute temperature

When the bulb of an ideal gas thermometer is immersed in water at the triple point, the height of the mercury column above the reference mark is 760 mm. The same thermometer is then immersed in a sample of unknown temperature, and the height of the mercury column moves to 1167 mm above the reference mark. What is the absolute temperature of the sample?

1 GETTING STARTED The height of the mercury column is linearly proportional to the absolute temperature. I therefore begin by making a graph like Figure 20.11, summarizing the information given in the problem (**Figure 20.12**). I denote the unknown absolute temperature by T and the height of the mercury at this unknown temperature by $h = 1167$ mm.

Figure 20.12



2 DEVISE PLAN Because the line giving height as a function of absolute temperature goes through the origin, its slope is equal to the ratio h_{tp}/T_{tp}. From my graph I see that the ratio h/T must have the same value: $h/T = h_{tp}/T_{tp}$. From this expression I can calculate the unknown temperature.

3 EXECUTE PLAN Solving for T, I get $T = T_{tp}h/h_{tp} = (273.16\ K)(1167\ mm)/(760\ mm) = 419\ K$. ✔

4 EVALUATE RESULT The absolute temperature I obtain is about 50% higher than T_{tp}, which is what I would expect, given that the mercury height is about 50% higher than that at the triple point.

In principle, the operation of the ideal gas thermometer as described above is independent of the gas used—we can use *any* gas we want. In practice, deviations from the ideal gas law complicate matters a bit, and the science and technology of gas thermometers are more complex than presented here. Nonetheless, gas thermometers serve as the standard reference for all other types of thermometers. Commercial thermometers typically give the temperature on a different scale (see the box "Temperature measurement scales" and **Table 20.1**). In thermodynamics, and in physics in general, it makes the most sense to use absolute temperatures. For simplicity we shall henceforth refer to the absolute temperature T expressed in kelvins simply as *temperature*.

In drawing Figure 20.11 we assumed that we can extrapolate the linear relationship between pressure and temperature from the ideal gas law down to zero. In practice, it is impossible for any gas to reach this temperature because as the temperature decreases, the gas condenses into a liquid. Of all gases, helium becomes liquid at the lowest temperature (4.22 K, which corresponds to −269 °C or −452 °F), and so its useful range extends below that of any other gas. Zero temperature ($T = 0$) is called **absolute zero.** Any system at absolute zero has zero entropy (that is, it has only one available basic state), and its thermal energy is a minimum. Absolute zero therefore is the limit of low temperature.

CONCEPTS

Table 20.1 Comparison of temperature scales (defining points in bold)

Temperature	Kelvin (K)	Celsius (°C)	Fahrenheit (°F)
Absolute zero	0	–273.15	–459.67
Water freezes (at a pressure of 101.325 kPa)	273.15	**0**	**32**
Triple point of water	**273.16**	0.01	32
Room temperature	293	20	68
Average human body temperature	310	37.0	98.6
Water boils (at a pressure of 101.325 kPa)	373.12	**100**	**212**
Titanium melts	1941	1668	3034
Surface of the Sun	5778	5505	9941

EXERCISE 20.2 Temperature conversions

Derive an expression for converting a temperature (*a*) from the Celsius scale to the Fahrenheit scale, (*b*) from the Kelvin scale to the Fahrenheit scale, and (*c*) from the Fahrenheit scale to the Kelvin scale.

SOLUTION (*a*) To obtain an expression for the temperature T_F measured in degrees Fahrenheit in terms of the temperature T_C measured in degrees Celsius, I use Eq. 2 from the "Temperature measurement scales" box, solving this expression for T_F:

$$T_F = \frac{9\ °F}{5\ °C} T_C + 32\ °F. \checkmark$$

(*b*) Substituting the expression for converting degrees Celsius to kelvins yields

$$T_F = \frac{9\ °F}{5\ °C}\left(\frac{1\ °C}{1\ K} T - 273.15\ °C\right) + 32\ °F. \checkmark$$

(*c*) I can use the expression for T_F I obtained in part *b*. Solving for T yields:

$$T = \left[\frac{1\ K}{1\ °C}\frac{5\ °C}{9\ °F}(T_F - 32\ °F)\right] + 273.15\ K. \checkmark$$

20.5 (*a*) During a certain day the outdoor temperature ranges between 60 °F and 80 °F. What is the corresponding range in degrees Celsius and in kelvins? (*b*) The filament temperature of an incandescent light bulb is about 2800 K. What is the corresponding temperature in degrees Celsius and in degrees Fahrenheit? (Assume only two significant digits in 2800 K.)

20.3 Heat capacity

The amount of energy it takes to raise the temperature of an object depends on the amount and kind of material in the object. Indeed, we know from experience that it takes longer to heat a large pot full of water than to heat the same pot containing only a cup of water. The larger quantity of water requires more energy, and therefore more gas needs to be burned.

We can characterize how much energy it takes to raise the temperature of an object by a scalar quantity called the object's **heat capacity.** The heat capacity of an object is equal to the amount of energy transferred thermally to the object divided by the resulting change in temperature.* The SI unit of heat capacity is J/K.

20.6 Which is greater in each pair: (*a*) the thermal energy of a pool of water or that of a glass of water at the same temperature, (*b*) the heat capacity of a pool of water or that of a glass of water at the same temperature, (*c*) the thermal energy of a cold glass of water or that of an equal amount of warm water?

Because heat capacity is an extensive quantity (it depends on the amount of material in an object), it makes sense to introduce a related intensive quantity so that we can compare the thermal properties of different materials. Traditionally, heat capacity is turned into an intensive quantity by dividing it by the mass of the object. The resulting quantity, called **specific heat capacity** and denoted by the lowercase letter c, is a measure of how much energy is required to raise the temperature of 1 kg of a certain material by 1 K. The adjective *specific* means "per unit mass." The SI unit of specific heat capacity is J/K · kg.

Table 20.2 on the next page gives the specific heat capacities of a variety of materials. As you can see, different materials have very different specific heat capacities. For example, it takes about ten times less energy to raise the temperature of a 1-kg block of copper by 1 K (385 J) than it takes to raise the temperature of an equal mass of water by 1 K (4181 J).

Why are these values so different from one another, and what determines the specific heat capacity of a material? Remember from Chapter 19 that when an object is in thermal equilibrium, its thermal energy is equally distributed among all the parts of the object that can store energy. Clearly, different materials have different abilities to store thermal energy. To understand these differences, let us first examine the heat capacity of an ideal gas.

*Strictly speaking this quantity is the average heat capacity, but because the heat capacity does not depend much on temperature, we can omit the word "average".

CONCEPTS

Table 20.2 Specific heat capacities

Material	c (J/K · kg)
alcohol	2436
aluminum	897
brass	385
carbon graphite	685
copper	385
glass	837
gold	130
helium	3116
ice (268 K)	2090
iron	449
lead	130
marble	860
mercury	140
nickel	443
platinum	133
silver	235
steam (373 K)	1556
titanium	527
water	4181
wood	1700
zinc	388

All values are reported at a temperature of 298 K (unless otherwise noted) and atmospheric pressure and with the gases held at constant volume.

Table 20.3 Heat capacities per particle in units of k_B

Class	Material	C_V/k_B
Monatomic gases	helium	1.50
	argon	1.50
	krypton	1.50
	xenon	1.51
	radon	1.51
Polyatomic gases	ammonia	3.37
	hydrogen	2.47
	nitrogen	2.50
Liquids	mercury	3.38
	water	8.97
Solids	aluminum	2.91
	copper	2.94
	gold	3.08
	iron	3.02
	silver	3.05
	titanium	3.04
	zinc	3.05

All values are reported at a temperature of 298 K and atmospheric pressure and with gases held at constant volume (as indicated by the subscript V on C_V).

In Chapter 19 we found that the thermal energy of a monatomic ideal gas is given by $E_{th} = \frac{3}{2}Nk_BT$ (Eq. 19.50). Each of the N ideal gas atoms thus has an average amount of thermal energy equal to $\frac{3}{2}k_BT$, where $k_B = 1.381 \times 10^{-23}$ J/K is the Boltzmann constant. We arrived at this result by considering the kinetic energy associated with the motion of an ideal gas atom along the x, y, and z directions (see Section 19.6). The motion along each axis is called a **degree of freedom.** Because there are three independent axes for translation, there are three "translational degrees of freedom." In thermal equilibrium, the equipartition of energy provides each translational degree of freedom of each atom with an equal share of the thermal energy—the **equipartition energy share** $\frac{1}{2}k_BT$—for a total of $\frac{3}{2}k_BT$ per atom.

If we raise the temperature of an ideal gas by 1 K, the thermal energy of the gas goes up by $\frac{3}{2}Nk_B$ (1 K), and that of each atom goes up by $\frac{3}{2}k_B$ (1 K). If this energy is transferred thermally to the gas, we expect the heat capacity per atom to be equal to $\frac{3}{2}k_B$.

Table 20.3 lists the heat capacities per particle, denoted by the uppercase letter C, in units of k_B for a number of substances. As you can see, it is indeed $\frac{3}{2}$ for monatomic gases, such as helium and argon, but deviates from this value for polyatomic gases. For hydrogen, for example, the value is closer to $\frac{5}{2}$ than to $\frac{3}{2}$, suggesting that a hydrogen molecule has two additional degrees of freedom that can store energy.

The reason for this discrepancy is that we treated each ideal gas particle as an atom that has only kinetic energy associated with translational motion along three independent axes. A dumbbell-shaped hydrogen molecule (Figure 20.13*a*), however, also has energy associated with rotational motion. The molecule can rotate about three independent axes (Figure 20.13*b*). Rotation about the z axis does not store any appreciable energy, because the rotational inertia about that axis is negligibly small.* So each hydrogen molecule has two additional "rotational degrees of freedom" for a total of five (three translational and two rotational). If each degree of freedom stores its equipartition share of $\frac{1}{2}k_BT$, we should obtain an energy per molecule of $\frac{5}{2}k_BT$. Consequently we expect a heat capacity per particle of $\frac{5}{2}k_B$, in good agreement with the experimental value of $2.47k_B$.

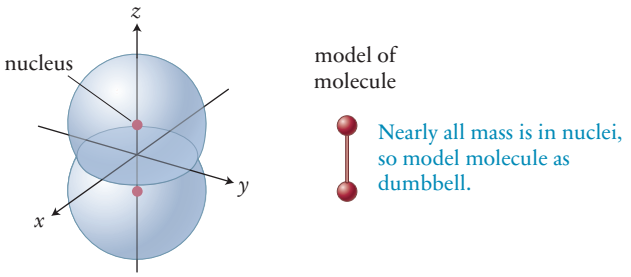
20.7 Consider the four-atom ammonia molecule illustrated in Figure 20.14. (*a*) Which of the three rotational degrees of freedom can store thermal energy? (*b*) Based on your answer to part *a*, what is the thermal energy per molecule? (*c*) What is the heat capacity per molecule? (*d*) Does your prediction in part *c* agree with the experimental value listed in Table 20.3?

*Although the atoms have a diameter of about 10^{-10} m and are separated by about the same distance, nearly all the mass is concentrated in the nucleus, which has a smaller diameter of about 10^{-15} m. Therefore the "bar" connecting the two atoms at the end of the H_2 dumbbell is about 10^5 times greater than the diameter of the atoms, and so the rotational inertia about the axis running through the length of the bar of the dumbbell is essentially zero (see Checkpoint 11.4).

CONCEPTS

Figure 20.13 A diatomic molecule has energy associated with rotation about two axes, as well as energy associated with translation along three axes.

(*a*) Diatomic molecule



(*b*) Rotational degrees of fredom for diatomic molecule

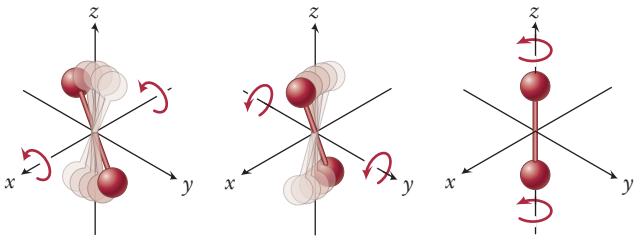


Figure 20.15 shows how, for a gas of diatomic hydrogen molecules, the heat capacity per particle in units of k_B varies with temperature. Notice how the heat capacity per particle in units of k_B is close to 2.5 (the value listed in Table 20.3) only around room temperature. As the temperature is lowered to about 50 K, the heat capacity per particle in units of k_B decreases to 1.5 $\left(\text{or } \frac{3}{2}\right)$. Why does hydrogen gas behave like a monatomic ideal gas at low temperatures? To ask the same question another way, why doesn't the rotational-degrees-of-freedom argument presented above for hydrogen hold at low temperatures? The experimental data suggest that, at low temperatures,

Figure 20.14 Ammonia molecule.

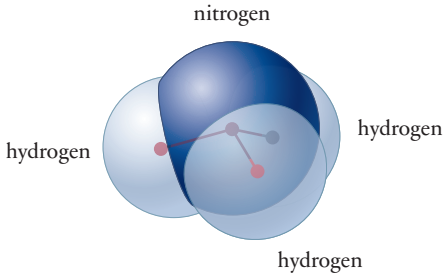
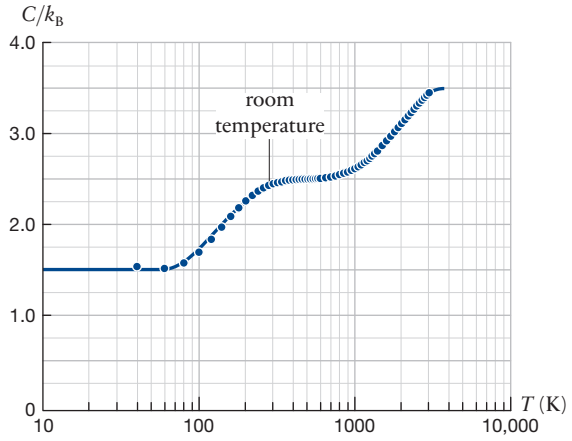


Figure 20.15 Temperature dependence of the heat capacity per particle C in units of k_B for hydrogen gas molecules.



the rotational motions of the molecules do not contribute to their thermal energy.

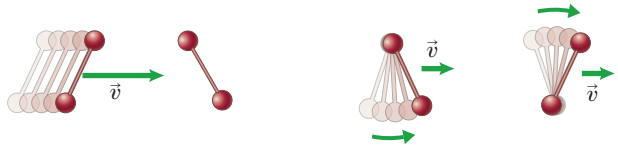
The reason for this behavior is that, according to quantum mechanics, vibrational and rotational energy comes in small fixed amounts, called quanta. Below about 100 K, the equipartition share $\frac{1}{2}k_BT$ is smaller than the quantum of energy needed to make a molecule rotate about one of the possible rotation axes. Therefore when one molecule collides with another, the amount of energy transferred during the collision is not great enough to set either molecule rotating (**Figure 20.16*a***). Above 100 K, the hydrogen molecules have sufficient kinetic energy to set each other spinning (Figure 20.16*b*). At room temperature (about 298 K),

Figure 20.16 Collisions between hydrogen molecules. (*a*) At low temperatures ($T < 100$ K), the molecules have insufficient kinetic energy to cause any changes in rotational motion. (*b*) At higher temperatures, molecules can be set into rotational motion.

(*a*)

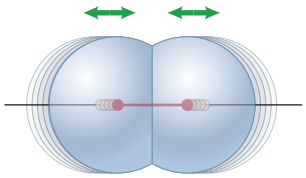


(*b*)



CONCEPTS

Figure 20.17 At temperatures above 1000 K, diatomic molecules begin to vibrate along the axis connecting the two atoms.



the value $\frac{1}{2}k_BT$ is equal to several rotational quanta, and so the equipartition of energy includes the two independent rotations of the hydrogen molecules, as described in the text preceding Checkpoint 20.7.

Figure 20.15 shows that at temperatures above 1000 K, the heat capacity per particle increases further. At such temperatures, the equipartition share becomes comparable to the quantum associated with another type of repetitive motion: vibration of the two hydrogen atoms along the axis of the molecule (**Figure 20.17**). Although there is only one axis for this type of motion in hydrogen, this vibrational motion contributes two degrees of freedom because it involves both potential and kinetic energy, each of which can store an amount of energy $\frac{1}{2}k_BT$. So the number of degrees of freedom for hydrogen is seven (three translational, two rotational, and two vibrational). If each degree of freedom contributes $\frac{1}{2}k_BT$, we expect a thermal energy per molecule of $\frac{7}{2}k_BT$. Indeed, experiments show that the heat capacity per hydrogen molecule approaches $\frac{7}{2}k_B$ around 3000 K. To summarize:

Only degrees of freedom for which $\frac{1}{2}k_BT$ is greater than the quantum of energy associated with that degree of freedom contribute to the heat capacity of a gas. In thermal equilibrium, each contributing degree of freedom of a particle in an ideal gas stores $\frac{1}{2}k_BT$ of thermal energy.

So, for a gas in which each particle has d contributing degrees of freedom, we expect the heat capacity per particle to be equal to $\frac{d}{2}k_B$.

CONCEPTS

EXAMPLE 20.3 Energy quanta for nitrogen

At room temperature, do collisions between nitrogen molecules have enough energy to set a nitrogen molecule rotating? Enough to set it vibrating? The quantum of energy associated with rotation is $E_{rot} = 4.0 \times 10^{-23}$ J and the quantum of energy associated with vibration is $E_{vib} = 4.7 \times 10^{-20}$ J.

1 GETTING STARTED At a given absolute temperature T each molecule stores $\frac{1}{2}k_BT$ of thermal energy per contributing degree of freedom. If the quantum of energy associated with a certain degree of freedom is greater than the energy stored per degree of freedom, then that degree of freedom cannot store any energy.

2 DEVISE PLAN To solve this problem, I need to compare the equipartition energy share at room temperature with the energy quanta associated with rotation and vibration. I can obtain the equipartition energy share at room temperature by substituting $T = 300$ K (room temperature) into $\frac{1}{2}k_BT$.

3 EXECUTE PLAN At 300 K the equipartition energy share is $\frac{1}{2}k_BT = \frac{1}{2}(1.38 \times 10^{-23}\ \text{J/K})(300\ \text{K}) = 2.07 \times 10^{-21}$ J. This value is greater than E_{rot} but smaller than E_{vib}, which means that at room temperature the nitrogen molecules rotate but do not vibrate. ✔

4 EVALUATE RESULT Table 20.3 confirms my conclusion: At room temperature the heat capacity per particle of nitrogen molecules is $2.50 \approx \frac{5}{2}$, indicating that rotations contribute to the storing of thermal energy but vibrations do not.

We can apply the same treatment to the vibrating atoms in a solid crystal. Consider, for example, the simple cubic arrangements of atoms shown in **Figure 20.18**. Each atom is "tied" in place by elastic bonds with neighboring atoms, so there are no translational or rotational degrees of freedom. However, each atom can vibrate along the three independent axes of the crystal. As we have seen, each vibration contributes two degrees of freedom, and so each atom should have six degrees of freedom, each storing $\frac{1}{2}k_BT$ of thermal energy. We therefore expect a heat capacity per atom of $3k_B$, which is in good agreement with the values for a number of simple crystalline solids (see Table 20.3).

Figure 20.19 shows the temperature dependence of the heat capacity per atom for a number of solids. Again we see that quantum effects affect the heat capacity. Below certain temperatures, the heat capacity of each solid decreases to zero because $\frac{1}{2}k_BT$ is smaller than the quantum of energy associated with the vibrations.

The interplay between the various degrees of freedom and the number of basic states available for the equipartition of energy explains the heat capacities for different materials over a broad range of temperatures. For a more accurate and complete description, however, we need *quantum mechanics*—a branch of physics that accounts for the discrete nature of energy and is beyond the scope of this text.

Figure 20.18 Each atom in a crystal can vibrate along three independent axes.

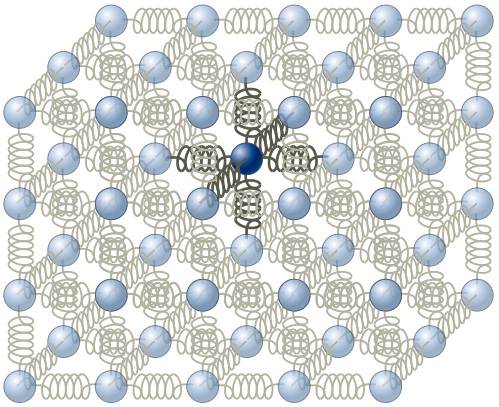
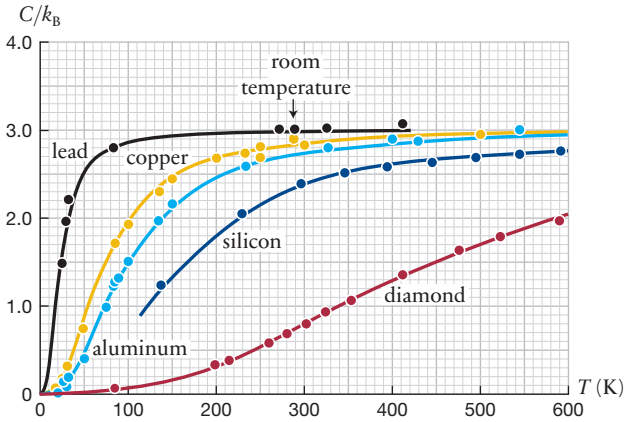


Figure 20.19 Temperature dependence of the heat capacity per atom for a number of solids.



In all of this discussion we have considered materials that do not undergo any phase transition from one state of matter to another. During a phase transition, adding energy thermally to a system does not change the temperature, but instead goes to changing the state of matter. (See the box "Phase transitions.")

20.8 Two identical volumes of gases—one monatomic and the other diatomic—are at the same pressure at room temperature. The temperature of each is raised by the same amount by thermally transferring energy to each of them. Is the energy transferred to the monatomic gas smaller than, equal to, or greater than the energy transferred to the diatomic gas?

20.4 *PV* diagrams and processes

We have seen that a system can exchange energy with its surroundings in two ways: by work done on it and by a thermal transfer of energy. The exchange of energy changes the state of the system. As we saw in Section 20.1, the system might get hotter, or expand, or the pressure might increase. In this section we examine the interplay between energy exchanges and changes in state for an ideal gas.

The equilibrium state of a gas can be completely specified by the four macroscopic measurable quantities we discussed in Chapter 19: the pressure P of the gas, its volume V, its temperature T, and the number of molecules N in that volume. Generally we shall consider samples that consist of a fixed number N of particles, so we can represent the equilibrium state of a gas by the triplet of values (P, V, T).

As we saw in Chapter 19, for an ideal gas the four macroscopic variables are related to one another by Eq. 19.51: $P = (N/V)k_BT$.* This means that if the number of particles in a gas is fixed, we need only two variables to completely specify the equilibrium state of the gas. If we choose these two variables to be P and V, then the equilibrium state of a fixed amount of ideal gas can be represented by a point on a *PV diagram* such as the ones shown in **Figure 20.20** on the next page.

*For a nonideal gas the four macroscopic variables are also related to one another, but the relationship is more complicated than that expressed by the ideal gas law.

Phase transitions

Most substances occur in several *states of matter* (also called *phases*): the solid state, the liquid state, and the gaseous state. (Plasma, a fourth state, contains ionized gas and free electrons.) H_2O, for example, exists in the solid state as ice, in the liquid state as water, and in the gaseous state as water vapor. The transition from one state of matter to another is called a *phase transition*. During a phase transition any energy transferred thermally does not change the temperature, but instead goes toward changing the state. At a given pressure, therefore, phase transitions occur at a fixed temperature and are accompanied by a change in volume.

For example, at normal atmospheric pressure ice changes to water at a temperature of 0 °C. To melt 1 kg of ice requires a thermal transfer of energy of 3.34×10^5 J. This energy is required to break the bonds that hold the water molecules together in ice.

Because the energy transferred thermally to a substance undergoing a phase transition is an extensive quantity, we introduce an intensive quantity by dividing the energy transferred thermally by the mass of the substance undergoing the transition. The resulting quantity, $L \equiv Q/m$, is called the *specific transformation energy* (traditionally often called *latent heat*, although the term *heat*, which is an energy and not a specific energy, is a misnomer). Because there are two phase transitions between the three states of matter, each substance has two specific transformation energies: one associated with the solid-liquid transition (*specific transformation energy for melting* L_m) and one with the liquid-gas transition (*specific transformation energy for vaporization* L_v).

Phase transitions are reversible. Energy must be added to a substance to make it undergo a transition from a lower- to a higher-temperature phase. In the other direction, energy must be removed from the system. For example, to change 1 kg of water at 100 °C to water vapor at 100 °C at normal atmospheric pressure requires thermally transferring 2.256×10^6 J to the water. When that vapor condenses back to water at 100 °C and normal atmospheric pressure, the same 2.256×10^6 J must be removed from the vapor.

Table 20.4 on the next page lists the specific transformation energies and associated phase transition temperatures for the melting and vaporization of a number of substances at normal atmospheric pressure.

CONCEPTS

Table 20.4 Phase transition temperatures and specific transformation energies for melting (m) and vaporization (v)

Substance	T_m (K)	L_m (10^3 J/kg)	T_v (K)	L_v (10^3 J/kg)
Helium	–	–	4.230	20.8
Hydrogen	13.84	58.6	20.37	449
Nitrogen	63.18	25.5	77.36	199
Oxygen	54.36	13.8	90.19	213
Ammonia	195	331	239.8	1369
Carbon tetrachloride	250.3	16	349.8	194
Ethanol	159	100.7	351	837
Water	273.15	334	373.12	2256
Mercury	234	11.8	630	272
Aluminum	933.6	395	2740	10,500
Lead	600.5	24.5	2023	871
Sulfur	392	38.1	717.75	326
Gold	1336.15	64.5	2933	1578
Copper	1356	134	2840	5069

We can read the corresponding pair of values (P, V) from the axes and determine the temperature from the ideal gas law (assuming we know N). Once we know the temperature, we can calculate the thermal energy and the entropy of the gas (see Eqs. 19.50 and 19.59 for monatomic ideal gases).

Any point on a *PV* diagram represents an equilibrium state in which the gas has a well-defined temperature, thermal energy, and entropy.

The ideal gas law implies that all equilibrium states (P, V) at a fixed temperature lie on a curve such that the product PV is constant or, to express this information another way, $P \propto 1/V$. These curves are called **isotherms.** Figure 20.20*a* shows a number of isotherms, each corresponding to a fixed temperature and therefore a fixed thermal energy of the gas. Isotherms that are higher on the PV diagram represent higher temperatures and greater thermal energy.

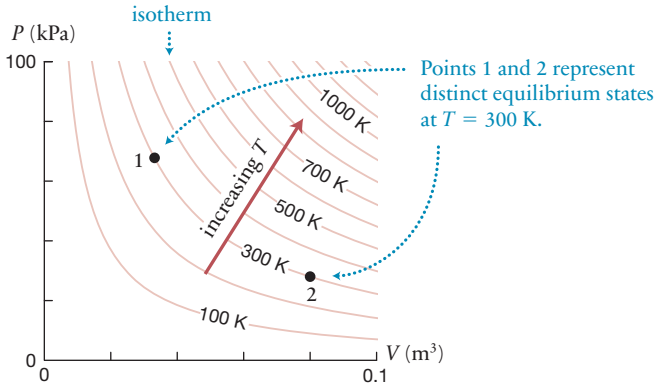
20.9 Consider the equilibrium states represented by points 1 and 2 in Figure 20.20*a*. Compare these quantities for an ideal gas in these two states: pressure P, volume V, temperature T, entropy S, and thermal energy E_{th}.

Another important family of curves associated with the PV diagram for an ideal gas is shown in Figure 20.20*b*. Each curve in this family connects all

CONCEPTS

Figure 20.20 Isotherms and isentropes in a PV diagram for an ideal gas. Points 1 and 2 represent two equilibrium states, both at $T = 300$ K. In general, on a PV diagram, the isentrope passing through a given point is steeper than the isotherm passing through that point.

(*a*) *PV* diagram showing isotherms (curves of constant temperature *T*) for an ideal gas



(*b*) Isentropes (curves of constant entropy *S*) superimposed on isotherms

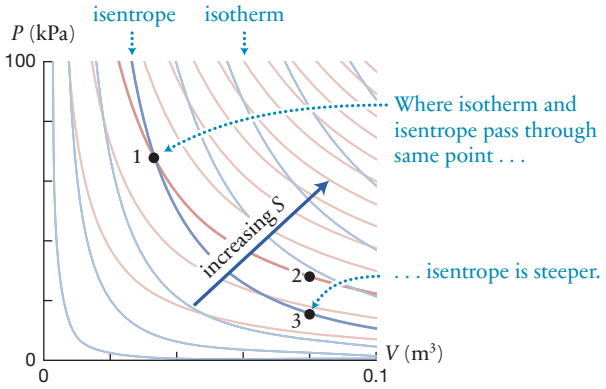
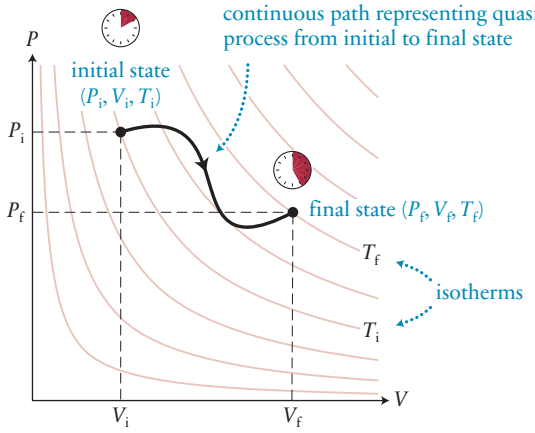


Figure 20.21 A path on a *PV* diagram represents a succession of equilibrium states that takes the system from some initial state to some final state. Such a transformation is called a quasistatic process.



points that correspond to equilibrium states that have the same entropy. Such curves are called **isentropes.**

An isentrope passing through a given point on a *PV* diagram is steeper than the isotherm passing through that point. To see why this is so, consider the isentrope and isotherm passing through point 1 in Figure 20.20*a*. As we saw in Checkpoint 20.9, the entropy of state 2 on the isotherm is greater than the entropy of state 1.* Equation 19.60, $S = N \ln(T^{3/2}V) + \text{constant}$, tells us that in order to return the entropy to its original value, the temperature must decrease as we go at constant volume from state 2 to state 3, which lies on the isentrope. So the isentrope passing through state 1 has to be steeper than the isotherm (Figure 20.20*b*).

Isentropes that are higher on the *PV* diagram represent greater entropies. You can see this by going from state 3 to state 2 in Figure 20.20*b*. The volume is the same for both states, but the fact that you move from one isotherm to a higher isotherm means that the temperature and thus the entropy (see Eq. 19.60) increase.

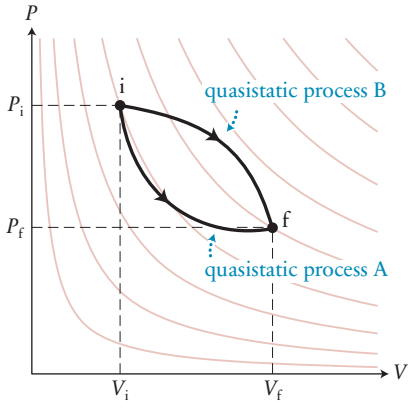
Let us now turn our attention to the transformation of a system from one equilibrium state to another. If the process is quasistatic, it is a succession of near-equilibrium states. Provided these states are very close to equilibrium, we can represent them by points on a *PV* diagram, and the succession of points then constitutes a path on the *PV* diagram, such as the one shown in **Figure 20.21**.

> *During a quasistatic process, the system remains near equilibrium at all instants. Such a process can be represented by a continuous path on a PV diagram.*

Processes that are not quasistatic cannot be represented on a *PV* diagram. For such processes we can only show the points that correspond to the initial and final equilibrium states.

*For simplicity, we shall omit the word *equilibrium* when referring to states presented by specific points on *PV* diagrams. For example, "state 1" stands for "the *equilibrium* state represented by point 1."

Figure 20.22 Two quasistatic processes that transform a gas from an initial state i to a final state f.



There are infinitely many ways to change the state of a gas. **Figure 20.22** shows two quasistatic processes that transform a gas from the same initial state i to the same final state f. Because the two processes have the same initial and final equilibrium states and because each state has a well-defined thermal energy and entropy, the change in these two quantities is independent of the process. For an ideal gas whose center of mass is at rest, there are no changes in the kinetic, potential, or source energy. Thus the change in the energy of the gas is equal to the change in the thermal energy, and we can state:

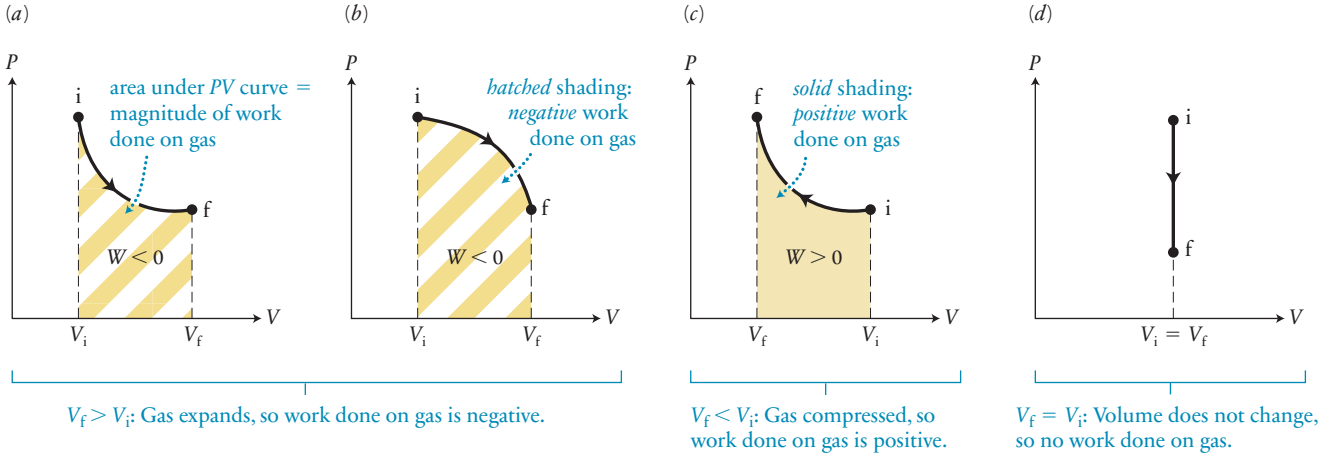
> *When an ideal gas is transformed from one equilibrium state to another, the changes in the thermal energy and entropy of the gas are independent of the process.*

Note that this statement holds even for non-quasistatic, irreversible processes between equilibrium states. The changes in thermal energy and entropy depend on only the initial and final states—what happens in between does not matter. If the change in thermal energy is independent of the process, then the energy transferred to the system $(W + Q)$ must also be independent of the process. Are W and Q each independent of the process, too?

As we shall see in Section 20.5, the magnitude of the work done on a gas during any quasistatic process carried out on the gas is given by the area under the path that represents the process in a *PV* diagram. **Figure 20.23** on the next page shows a number of quasistatic processes, with the area under the path representing each process shaded; the larger the shaded area, the greater the magnitude of the work done on the gas. If the path is in the direction of increasing volume (Figures 20.23*a* and *b*), the work done on the gas is negative. If the path is in the direction of decreasing volume (Figure 20.23*c*), the work done on the gas is positive. For a vertical path (Figure 20.23*d*), the work is zero.

CONCEPTS

Figure 20.23 The magnitude of the work done on a gas during a quasistatic process is equal to the area under the path on the PV diagram representing that process.



Comparing the sizes of the hatched areas in parts *a* and *b* of Figure 20.23 tells us that the amount of work done on the gas is different in the two graphs. These graphs represent processes A and B of Figure 20.22, telling us that the magnitude of the work needed to carry out process B is greater than the work needed to carry out process A. Unlike changes in thermal energy that occur during a process, the amount of work done does depend on the path taken.

20.10 (*a*) For which of the two quasistatic processes shown in Figure 20.22, if either, is the change in entropy greater? (*b*) Is the energy transferred thermally to the gas the same along both paths?

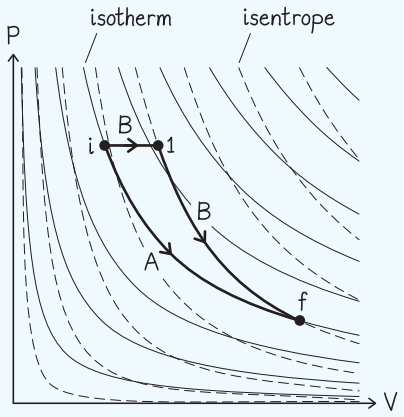
As Checkpoint 20.10 makes clear, the quantity of energy Q transferred thermally to the system, like the work W done on it, depends on the path taken.

CONCEPTS

EXAMPLE 20.4 Different process, same change?

In quasistatic process A of **Figure 20.24**, an ideal gas is brought from an initial to a final state while its temperature is held fixed. In quasistatic process B, an identical gas is brought in two steps from the same initial state as in process A to the same final state. In the first step, the gas expands at constant pressure from the initial state to an intermediate state 1, and in the second step, the gas is brought adiabatically to the final state. Is the quantity of energy transferred thermally to the gas during process B greater than, equal to, or smaller than during process A?

Figure 20.24 States i and f are on the same isotherm.



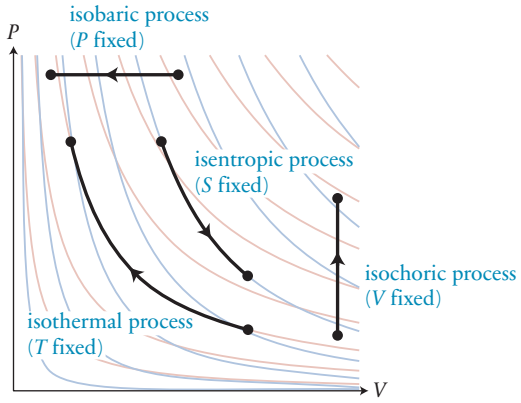
1 GETTING STARTED During both processes, the gas expands, and so the work done on the gas is negative in both cases. The quantity of energy transferred thermally depends on both the work done on the gas and the change in the thermal energy of the gas.

2 DEVISE PLAN The initial and final states lie on the same isotherm, which means the temperature of the gas is the same in the two states. Therefore the thermal energy of the gas is the same in the two states, and so the change in thermal energy is zero for both processes. This means that the quantity of energy transferred thermally to the gas is the negative of the work done on the gas.

3 EXECUTE PLAN The area under the path for process B in Figure 20.24 is larger than the area under the path for process A, and so the magnitude of the work done on the gas is greater for process B. This means that the quantity of energy transferred thermally must be greater during process B. ✔

4 EVALUATE RESULT I am told that the second step in process B is adiabatic, which means that $Q = 0$. The problem therefore boils down to comparing the energy transferred thermally during process A with the energy transferred thermally during the first step of process B. During process A, the change in the thermal energy is zero, and so all the energy transferred thermally to the gas can be used by the gas to do work on its environment. That state 1 is on a higher isotherm than state i tells me that the temperature of the gas increases during the first step of process B. Therefore part of the energy transferred thermally is required to raise the thermal energy of the gas. It therefore makes sense that Q is greater for process B.

Figure 20.25 Four types of constrained quasistatic processes.



The quasistatic processes in Example 20.4 are examples of *constrained processes*: Their path on a *PV* diagram is determined by a constraint placed on the system.

Figure 20.25 shows four important constrained quasistatic processes in ideal gases that we shall study in the second part of this chapter. When the pressure in the gas is held fixed during a quasistatic process, the path in the *PV* diagram is a horizontal line at the fixed pressure *P*; such a process is called isobaric (Greek for "same weight"). When the volume is held fixed, the path is a vertical line, and the process is called isochoric ("same volume"). Because the volume doesn't change, no work is done on (or by) the gas in an isochoric process (Figure 20.23*d*).

When the temperature is held fixed, the path lies along an isotherm, and the process is called isothermal ("same temperature"). As we saw in Chapter 19, the thermal energy of an ideal gas is proportional to the temperature of the gas, and so in an isothermal process, the thermal energy does not change.

Table 20.5 Constrained quasistatic processes

Process	Constraint	Consequence
isochoric	V fixed	$W = 0$
isentropic (adiabatic)	S fixed	$Q = 0$
isobaric	P fixed	
isothermal	T fixed	$\Delta E_{\text{th}} = 0$

When the entropy is constant during a process, the path lies along an isentrope and the process is called isentropic ("same entropy"). As we saw in Section 20.1, the entropy of a system remains constant during quasistatic adiabatic processes, and therefore these processes are isentropic. Paths along isentropes thus represent processes for which no energy is transferred thermally to the system ($Q = 0$). Isentropes are sometimes also called *adiabats*, and isentropic processes are also called *adiabatic processes*.*

Table 20.5 summarizes these constrained quasistatic processes and the consequences of the constraints imposed.

20.11 For each process in Figure 20.25, determine whether the following quantities are positive, negative, or zero: W, Q, ΔE_{th}, ΔS.

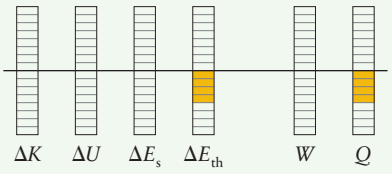
*Note, however, that an adiabatic process need not be isentropic. Only a *quasistatic* adiabatic process is isentropic.

CONCEPTS

Self-quiz

1. **Figure 20.26** shows the energy changes for a certain system over a short time interval. Describe what is happening to the system. From the diagram, what can you conclude about the temperature of the environment at the beginning of the time interval?

Figure 20.26



2. On a really hot day, which temperature scale (°F or °C) registers the greater number? On a really cold day, which temperature scale registers the smaller number? At what temperature do the two scales register the same number?
3. Which material in Table 20.2 has the greatest specific heat capacity? The smallest? Imagine that objects made of these materials are placed on your hand. The objects have the same mass and are initially at the same temperature. Each object equilibrates thermally with your hand, which is at a higher temperature. Describe any differences in the equilibration process.
4. What is the approximate thermal energy of a molecule that has three degrees of translational freedom, two degrees of rotational freedom, and two degrees of vibrational freedom? The molecule is in a gas at temperature T.
5. Which types of processes are represented in Figure 20.24? For each process, determine whether the following quantities are positive, negative, or zero: ΔT, ΔV, ΔP.

Answers

1. The thermal energy of the system is decreasing ($\Delta E_{\text{th}} < 0$) because energy is leaving the system ($Q < 0$) and entering the environment. That energy leaves the system means that initially the temperature of the environment must be lower than the temperature of the system.
2. The Fahrenheit scale registers the greater number on a really hot day. If the Fahrenheit temperature is 110 °F, the Celsius temperature is 43.3 °C. The Celsius scale registers the smaller number on a really cold day. If the Fahrenheit temperature is −20 °F, the Celsius temperature is −29 °C. The scales read the same temperature at −40 °C = −40 °F.
3. Water is the highest, gold and lead the lowest. Because water has such a high specific heat capacity, a large amount of energy must leave your hand and enter the water as your hand and the water equilibrate, much more than the gold or lead would. This, in part, is why it is so hard to keep warm if you are wet and why gold jewelry so quickly comes to the same temperature as your skin.
4. With seven degrees of freedom and an equipartition energy share of $\frac{1}{2}k_{\text{B}}T$, the molecule has roughly $\frac{7}{2}k_{\text{B}}T$ of thermal energy.
5. Process A in Figure 20.24 is isothermal because it follows an isotherm. $\Delta T = 0$ because the process is isothermal, $\Delta V > 0$ because $V_{\text{f}} > V_{\text{i}}$, and $\Delta P < 0$ because $P_{\text{f}} < P_{\text{i}}$. The first step of process B is isobaric because the pressure is constant. $\Delta T > 0$ because T_{i} is on a lower isotherm than T_1, $\Delta V > 0$ because $V_1 > V_{\text{i}}$, and $\Delta P = 0$ because $P_1 = P_{\text{i}}$. The second step of process B is isentropic because it follows an isentrope. $\Delta T < 0$ because T_{f} is on a lower isotherm than T_1, $\Delta V > 0$ because $V_{\text{f}} > V_1$, and $\Delta P < 0$ because $P_{\text{f}} < P_1$.

20.5 Change in energy and work

If we take into account both the transfer of mechanical energy (work) W and the energy transferred thermally Q, the energy law (Eq. 9.1) becomes

$$\Delta E = W + Q. \tag{20.1}$$

Remember that this law is a direct consequence of conservation of energy: The only way the energy of a system can change is by a transfer of some form of energy to or from the surroundings.

Let's first consider only situations where the system is an ideal gas and where the translational kinetic energy, potential energy, and source energy of the gas remain constant ($\Delta K = \Delta U = \Delta E_s = 0$). The energy law then reduces to

$$\Delta E_{th} = W + Q \quad \text{(ideal gas)}. \tag{20.2}$$

Equation 19.50 tells us that the thermal energy of a monatomic ideal gas consisting of N atoms is

$$E_{th} = \tfrac{3}{2} N k_B T \quad \text{(monatomic ideal gas)}. \tag{20.3}$$

As we saw in Section 20.3, at thermal equilibrium each degree of freedom of an ideal gas particle contributes $\frac{1}{2} k_B T$ to the particle's thermal energy. For a monatomic ideal gas, each atom has three degrees of freedom, yielding the factor 3 in Eq. 20.3. Generalizing this result for an ideal gas with d contributing degrees of freedom, we have

$$E_{th} = \frac{d}{2} N k_B T \quad \text{(ideal gas)}. \tag{20.4}$$



20.12 What is the thermal energy associated with a gas sample that contains N hydrogen molecules at temperature $T = 300$ K?

When the temperature of an ideal gas sample that contains a fixed number N of particles changes, Eq. 20.4 tells us that the thermal energy of the sample changes by an amount

$$\Delta E_{th} = \frac{d}{2} N k_B \Delta T. \tag{20.5}$$

As we noted in Section 20.4, this change in thermal energy depends on only the initial and final equilibrium states and is independent of the path taken from initial state to final state.

To calculate the work done on a gas, consider the situation shown in **Figure 20.27**: A piston with surface area A quasistatically compresses a gas in a cylinder. The x component of the force exerted by the piston on the gas is F^c_{pgx}. The work done by this external force on the gas is (Eq. 9.22)

$$W = \int_{x_i}^{x_f} F^c_{pgx}\, dx. \tag{20.6}$$

The magnitude of the force exerted by the piston on the gas is related to the gas pressure P by Eq. 19.43:

$$P = \frac{F^c_{pg}}{A} \tag{20.7}$$

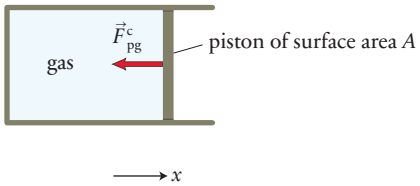


Figure 20.27 A piston that has a surface area A exerts a force on a gas, compressing it.

or $F^c_{pg} = PA$. If we take the positive x axis to the right as in Figure 20.27, we have $F^c_{pg,x} = -F^c_{pg} = -PA$. When the piston moves a distance dx, the volume of the gas changes by an amount $dV = A\,dx$, and so the work done on the gas is

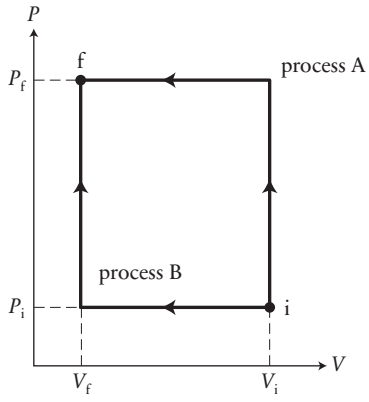
$$W = -\int_{x_i}^{x_f} PA\,dx = -\int_{V_i}^{V_f} P\,dV. \tag{20.8}$$

The integral to the right of the minus sign in this expression represents the area under the path taken in the PV diagram, as we discussed in Section 20.4 (see Figure 20.23). If the process is in the direction of increasing volume (expansion), the integral in Eq. 20.8 is positive, making the work done on the gas negative (Figures 20.23*a* and *b*). If the process is in the direction of decreasing volume (compression), the work done on the gas is positive (Figure 20.23*c*).

When the pressure is constant (that is, during a quasistatic isobaric process), we can pull the pressure out of the integral, and the work done on the gas is

$$W = -P\int_{V_i}^{V_f} dV = -P(V_f - V_i) = -P\Delta V \quad \text{(quasistatic isobaric process)}. \tag{20.9}$$

Figure 20.28 Checkpoint 20.13 and Example 20.5.



20.13 Which quasistatic process from the initial state i to the final state f in **Figure 20.28** requires more work done on the gas?

From Checkpoint 20.13 we see that the work done on a gas depends on the manner in which the gas is transformed from its initial state to its final state.

EXAMPLE 20.5 Same change, different process

In Figure 20.28, an ideal gas is brought from an initial state i to a final state f by two processes. The initial volume and pressure are 0.50 m^3 and 100 kPa, and the final values for these variables are 0.10 m^3 and 500 kPa. How much energy is transferred thermally to the gas during (*a*) process A and (*b*) process B?

❶ **GETTING STARTED** Both processes are represented by a continuous path on a PV diagram and so are quasistatic. From the graph and the data given, I know that $V_i = 0.50$ m^3, $V_f = 0.10$ m^3, $P_i = 100$ kPa, and $P_f = 500$ kPa.

❷ **DEVISE PLAN** To calculate the quantity of energy transferred thermally, I can use the energy law as given in Eq. 20.2, $\Delta E_{th} = W + Q$, which gives me $Q = \Delta E_{th} - W$. From Checkpoint 20.13, I know that work is done on the gas only during the constant-pressure legs. To calculate W during these legs, I can use Eq. 20.9. The change in the thermal energy ΔE_{th} is given by Eq. 20.5, but I do not know d, N, or ΔT. The ideal gas law, however, relates pressure and volume to temperature: $PV = Nk_BT$ (Eq. 19.51). I note that the product PV is the same for the initial and final states: $P_iV_i = (100\text{ kPa})(0.50\text{ m}^3) = 50\text{ kPa}\cdot\text{m}^3$ and $P_fV_f = (500\text{ kPa})(0.10\text{ m}^3) = 50\text{ kPa}\cdot\text{m}^3$. This tells me that the initial and final states lie on an isotherm, so $\Delta T = 0$. Therefore Eq. 20.5 becomes $\Delta E_{th} = 0$, so Eq. 20.2 yields $Q = -W$.

❸ **EXECUTE PLAN** The change in volume for both processes is $\Delta V = V_f - V_i = (0.10\text{ m}^3) - (0.50\text{ m}^3) = -0.40\text{ m}^3$.

(*a*) From Eq. 20.9, the work done on the gas for process A is $W = -P\Delta V = -(500\text{ kPa})(-0.40\text{ m}^3) = 2.0\times10^2$ kJ. Therefore $Q_A = -2.0\times10^2$ kJ. ✔

(*b*) $W = -P\Delta V = -(100\text{ kPa})(-0.40\text{ m}^3) = 40$ kJ, which leads to $Q_B = -40$ kJ. ✔

❹ **EVALUATE RESULT** The work done on the gas during process A is five times greater than the work done during process B. This is not surprising because in A the compression from V_i to V_f is done at a pressure that is five times greater than the pressure at which the compression is done in B. Because the thermal energy of the gas does not change, any energy acquired in the form of work done on the gas has to leave as energy transferred thermally out of the gas. Both Q_A and Q_B should therefore be negative, and Q_A should have a greater negative value than Q_B, which is what I found.

As we have seen, the change in thermal energy for an ideal gas undergoing some quasistatic process is independent of the process and is given by Eq. 20.5. Checkpoint 20.13 and Example 20.5, however, show that the work done on the gas and the energy transferred thermally depend on the process.

20.14 Consider the three processes shown in **Figure 20.29**. Process B is isothermal. (*a*) For which of the three processes is the energy transferred thermally greatest? (*b*) For which of the three processes is the energy transferred thermally smallest?

20.6 Isochoric and isentropic ideal gas processes

Let us now take a deeper look at what the energy law tells us about the transfers of mechanical and thermal energy that occur during constrained quasistatic processes. We begin by examining isochoric (constant volume) and isentropic (constant entropy) quasistatic processes in this section and then isobaric (constant pressure) and isothermal (constant temperature) quasistatic processes in the next section.

An example of an isochoric quasistatic process is shown in **Figure 20.30**. An ideal gas sample consisting of N particles enclosed in a fixed volume is heated. The energy diagram for the volume of ideal gas during this process is shown in Figure 20.30*b*. Because the gas is heated, its temperature and therefore its thermal energy increase. Because the volume is fixed, no work is done on the gas. The amount Q of energy transferred thermally to the gas raises the thermal energy of the gas.

Figure 20.30*c* shows the corresponding PV diagram. The ideal gas law, $P = Nk_B T/V$ (Eq. 19.51), tells us that if the volume is kept fixed, the pressure in the gas increases as the temperature increases. Because the volume is fixed and the pressure increases, the path that represents the heating process is straight up. The arrow representing Q has solid shading because energy is transferred to the gas ($Q > 0$).

Because $V_i = V_f$, the integral in Eq. 20.8 is zero, and so the work done on the gas is zero:

$$W = 0 \quad \text{(isochoric process)}. \tag{20.10}$$

The energy law for an ideal gas, $\Delta E_{th} = W + Q$, then reduces to

$$\Delta E_{th} = Q \quad \text{(isochoric process)}. \tag{20.11}$$

Substituting the expression given for ΔE_{th} in Eq. 20.5, we see that

$$Q = \frac{d}{2} N k_B \Delta T \quad \text{(isochoric process)}. \tag{20.12}$$

In Section 20.3, we introduced the quantity **heat capacity per particle**. Mathematically this quantity is defined as the energy transferred thermally divided by the number of particles and by the temperature change:

$$C_V \equiv \frac{Q}{N\Delta T} \quad \text{(constant volume)}, \tag{20.13}$$

where here we add a subscript V to remind ourselves that this result is obtained at constant volume. Substituting Eq. 20.12 into this definition for C_V yields

$$C_V = \frac{d}{2} k_B \quad \text{(constant volume)}. \tag{20.14}$$

Figure 20.29 Checkpoint 20.14.

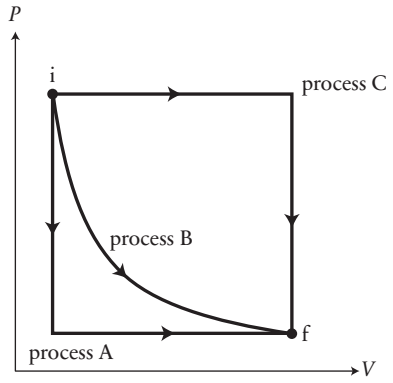
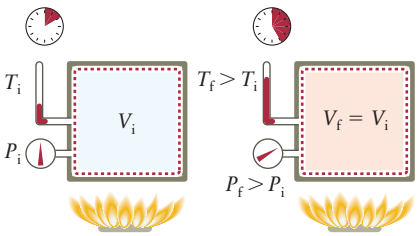
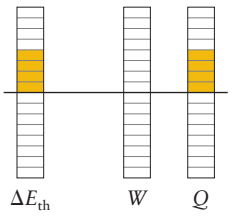


Figure 20.30 An example of an isochoric quasistatic process.

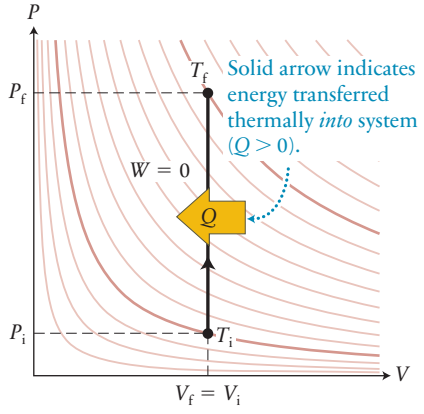
(*a*) Fixed volume of gas is heated, raising T and P



(*b*) Energy diagram for gas



(*c*) PV diagram showing the isochoric process



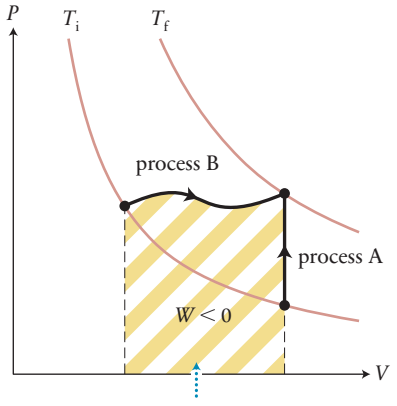
Note that this result predicts that the heat capacity per particle is determined by the number of degrees of freedom, in agreement with the experimental results we discussed in Section 20.3.

Equation 20.14 allows us to rewrite Eq. 20.5 in the form

$$\Delta E_{\text{th}} = NC_V\Delta T \quad \text{(any ideal gas process)}. \tag{20.15}$$

Even though we used Eq. 20.5 to arrive at Eq. 20.15 for ΔE_{th} under isochoric conditions, Eq. 20.15 holds for all ideal gas processes, not just isochoric ones. To see why, consider two samples of N ideal gas particles undergoing the two processes in **Figure 20.31**. Both samples are initially at temperature T_i and end at temperature T_f. Process A is isochoric; process B is some arbitrary process. Because the temperature change is the same for the two processes, Eq. 20.5 tells us the thermal energy change must also be the same. This means we can use Eq. 20.15 to calculate ΔE_{th} for any process, given the change in temperature ΔT in the process and knowing C_V for the gas. In contrast to Eq. 20.5, which expresses the thermal-energy change in terms of the number of degrees of freedom of each ideal gas particle, Eq. 20.15 gives the thermal-energy change in terms of macroscopically measured quantities.

Figure 20.31 Two ideal gas processes. Process A is isochoric; process B is not.



Hatched shading: Work done on gas in process B is *negative*.

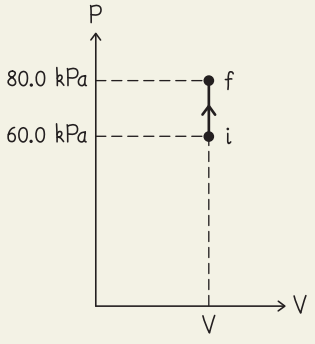
20.15 Is Eq. 20.15 valid for an isothermal process ($\Delta T = 0$)?

EXAMPLE 20.6 Energy transferred thermally

A sample of hydrogen gas that contains 1.00 mol of H_2 molecules at a temperature of 300 K and a pressure of 60.0 kPa undergoes an isochoric process. The final pressure is 80.0 kPa. How much energy is transferred thermally to the gas?

1 GETTING STARTED I begin by making a PV diagram of the process, treating the hydrogen as an ideal gas (**Figure 20.32**). Because the process is isochoric, I represent it by a vertical line. Because $P_f > P_i$, I indicate that the path representing the process points upward.

Figure 20.32



2 DEVISE PLAN The quantity of energy transferred thermally during an isochoric process is given by Eq. 20.12. To use this equation, I need to know N, d, and ΔT. I know from Avogadro's number that my 1.00-mol sample contains $N = 6.02 \times 10^{23}$ molecules, and I can obtain d from Eq. 20.14 if I use Table 20.3 to get the value of C_V/k_B for H_2 molecules. For ΔT, I am given the initial temperature but not the final temperature. I know the initial and final pressures, however, and because the volume is fixed in this isochoric process, I can use the ideal gas law to determine the final temperature and so have a value for ΔT.

3 EXECUTE PLAN From Table 20.3 I see that, for H_2 molecules, $C_V/k_B = 2.47$ at room temperature (about 300 K). From Figure 20.15 I see that C_V/k_B stays at that value through 700 K, and so, from Eq. 20.14, I know that $d/2 = C_V/k_B = 2.47$.

Next I rewrite the ideal gas law (Eq. 19.51) in the form $V = Nk_BT/P$. Because V, N, and k_B are constant, I know that T/P is constant, and so $T_i/P_i = T_f/P_f$. Solving for T_f, I get

$$T_f = \frac{T_iP_f}{P_i} = \frac{(300\text{ K})(80.0\text{ kPa})}{60.0\text{ kPa}} = 400\text{ K},$$

so $\Delta T = 400\text{ K} - 300\text{ K} = 100\text{ K}$. Substituting my values for N, d, and ΔT in Eq. 20.12 gives me

$$Q = \frac{d}{2}Nk_B\Delta T = 2.47(6.02\times10^{23})(1.38\times10^{-23}\text{ J/K})(100\text{ K})$$
$$= 2.05\times10^{3}\text{ J}. \checkmark$$

4 EVALUATE RESULT Because the temperature of the gas increases, its thermal energy must increase. The process is isochoric, so no work is done on the gas. Therefore the value of the quantity of energy transferred thermally must be positive, as my result is.

The quantity of energy transferred is significant—approximately equal to the kinetic energy of a 1-kg object moving at about 65 m/s—but I know that thermal energies tend to be large relative to kinetic energies. (See, for example, the box "Coherent versus incoherent energy" in Chapter 7.)

QUANTITATIVE TOOLS

Next we consider a quasistatic isentropic process, a process for which the entropy remains unchanged:

$$\Delta S = 0 \quad \text{(isentropic process)}. \tag{20.16}$$

As we saw in Section 20.4, no energy is transferred thermally in an isentropic quasistatic process:

$$Q = 0 \quad \text{(isentropic quasistatic process)}. \tag{20.17}$$

An example of an isentropic quasistatic process is shown in **Figure 20.33*a***. An ideal gas sample consisting of N particles is compressed very slowly by grains of sand dropped on the piston to ensure that the process is quasistatic. The gas is thermally insulated so that $Q = 0$. The increasing force exerted on the piston increases the pressure in the gas and decreases its volume. The corresponding energy diagram for the volume of ideal gas is shown in Figure 20.33*b*. The work done by the piston on the gas is positive, increasing the thermal energy of the gas. Because $Q = 0$, Eq. 20.2 becomes

$$\Delta E_{\text{th}} = W \quad \text{(isentropic quasistatic process)}. \tag{20.18}$$

The increase in thermal energy means that the temperature of the gas goes up even though the gas is not being heated. During an isentropic compression, work done on the gas raises the temperature of the gas. You may have noticed this effect while using a bicycle pump: If you pump vigorously, the air in the pump—and therefore the pump and the tire tube—becomes noticeably hotter.

Figure 20.33*c* shows the PV diagram for the isentropic process. The shaded area under the path is the work done by the piston on the gas. As the solid shading indicates, this work is positive ($W > 0$). Because the entropy of the gas doesn't change when $Q = 0$, the path in Figure 20.33*c* lies along an isentrope, and because the gas is compressed, the final volume is smaller than the initial volume. As you can see from the PV diagram, the path crosses the isotherms and ends on an isotherm at a higher temperature, indicating that the process raises the temperature of the gas, as we saw before.

If we substitute Eq. 20.15 into Eq. 20.18, we obtain an expression that relates the work done on the gas W to the temperature change:

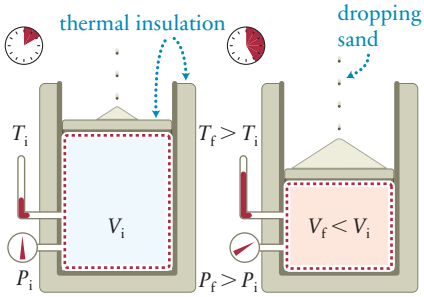
$$\Delta E_{\text{th}} = W = NC_V\Delta T \quad \text{(isentropic process)}. \tag{20.19}$$



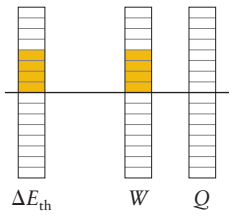
20.16 Is the product PV in the final state in Figure 20.33*c* greater than, equal to, or smaller than the product PV in the initial state?

Figure 20.33 An example of an isentropic quasistatic process.

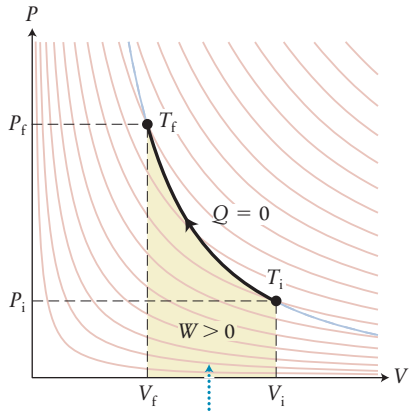
(*a*) Thermally insulated volume of ideal gas is compressed slowly



(*b*) Energy diagram for gas



(*c*) PV diagram showing the isentropic process



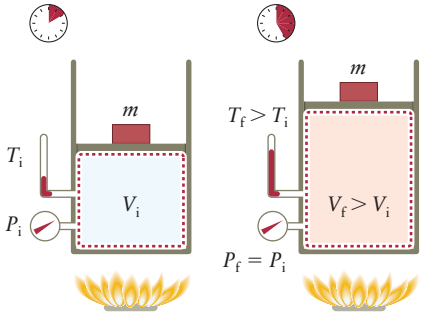
Solid shading: work done on gas is *positive*.

20.7 Isobaric and isothermal ideal gas processes

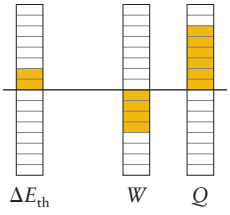
Next we turn to an isobaric (constant pressure) quasistatic process. An example is shown in **Figure 20.34*a*** on the next page: An ideal gas sample consisting of N particles is heated while it is being held at constant pressure by a piston that exerts a constant force on the gas. If the piston were held in place, the heating would raise the temperature of the gas and, according to the ideal gas law, the pressure would go up. Because the piston is free to move, however, the pressure stays constant and the gas expands instead. Therefore the gas does work on its environment. The energy diagram for the volume of the ideal gas during this process is

Figure 20.34 An example of an isobaric quasistatic process.

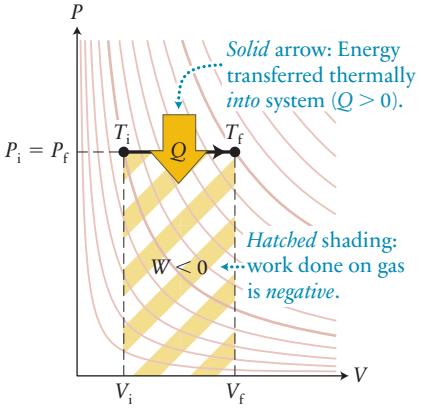
(*a*) Ideal gas heated at constant pressure



(*b*) Energy diagram for gas



(*c*) *PV* diagram showing the isobaric process



shown in Figure 20.34*b*. The heating transfers energy to the gas ($Q > 0$), part of which goes into raising the temperature (and therefore the thermal energy of the gas: $\Delta E_{th} > 0$), and part of which is used by the gas to do work on its environment ($W < 0$).

Figure 20.34*c* shows the *PV* diagram for the process. Because the pressure is constant, the path lies on a horizontal line. Because the gas expands, the direction of the path is to the right, toward increasing volume. The shaded area under the path is the negative of the work done on the gas ($W < 0$). (Remember our convention: A hatched area signifies that the work done *on the gas* is negative.) We draw an arrow representing Q with solid shading to represent the fact that we transfer energy thermally to the gas ($Q > 0$).

Because this process is at constant pressure, we can define heat capacity per particle at constant pressure in analogy to Eq. 20.13:

$$C_P \equiv \frac{Q}{N\Delta t} \quad \text{(constant pressure)}, \tag{20.20}$$

and so in an isobaric process the energy transferred thermally is

$$Q = NC_P\Delta T \quad \text{(isobaric process)}. \tag{20.21}$$

From the ideal gas law (Eq. 19.51), $PV = Nk_BT$, we have for the initial and final states $PV_i = Nk_BT_i$ and $PV_f = Nk_BT_f$. The change in volume is therefore $\Delta V = V_f - V_i = Nk_B\Delta T/P$, which means the work done on the gas is, from Eq. 20.9,

$$W = -P\Delta V = -Nk_B\Delta T \quad \text{(isobaric process)}. \tag{20.22}$$

Now we have all three terms in the energy law for an ideal gas, $\Delta E_{th} = W + Q$. Equation 20.15 gives the change in the thermal energy ΔE_{th}, Eq. 20.22 gives the work done on the gas W, and Eq. 20.21 gives the energy Q transferred thermally into the gas. We can use this information to obtain an expression for C_P for ideal gases. Substituting Eqs. 20.15, 20.21, and 20.22 into the energy law for an ideal gas, we obtain

$$NC_V\Delta T = -Nk_B\Delta T + NC_P\Delta T \quad \text{(ideal gas)}. \tag{20.23}$$

Dividing by $N\Delta T$, we see that for an ideal gas, the heat capacities per particle are related by the expression

$$C_P = C_V + k_B \quad \text{(ideal gas)} \tag{20.24}$$

or, substituting from Eq. 20.14 for C_V,

$$C_P = \frac{d}{2}k_B + k_B = \left(\frac{d}{2} + 1\right)k_B \quad \text{(ideal gas)}. \tag{20.25}$$

Using Eqs. 20.14 and 20.25, we can now write the **heat capacity ratio** as

$$\gamma \equiv \frac{C_P}{C_V} = \frac{\left(\frac{d}{2} + 1\right)k_B}{\frac{d}{2}k_B} = 1 + \frac{2}{d} \quad \text{(ideal gas)}. \tag{20.26}$$

For a monatomic ideal gas, for which $d = 3$, the heat capacity ratio is $\gamma = \frac{5}{3}$.

QUANTITATIVE TOOLS

EXAMPLE 20.7 Heat capacity ratio

Experimentally one obtains $C_P = 4.844 \times 10^{-23}$ J/K for nitrogen gas at room temperature. (*a*) Calculate the heat capacity ratio γ for a gas of nitrogen molecules, N_2, at room temperature. (*b*) Which degrees of freedom of the N_2 molecules contribute to the heat capacity?

❶ **GETTING STARTED** I am given the room-temperature C_P value for nitrogen gas and asked to calculate γ and then to determine which of the degrees of freedom of the N_2 molecules contribute to the heat capacity. I assume the gas behaves ideally so that I can apply the results of this section.

❷ **DEVISE PLAN** The heat capacity ratio γ is given by Eq. 20.26. To calculate this ratio I need to know d, the number of degrees of freedom of the N_2 molecules that contribute to storing energy at room temperature. Equation 20.25 relates d to C_P, which is given. Once I determine d using this relationship, I can calculate γ and establish which degrees of freedom contribute to the heat capacity.

❸ **EXECUTE PLAN** (*a*) From Eq. 20.25, I get $\frac{d}{2} + 1 = C_P/k_B$, or $d = 2(C_P/k_B - 1)$. Substituting the value given for C_P, I get

$$d = 2\left(\frac{4.844 \times 10^{-23}\ \text{J/K}}{1.381 \times 10^{-23}\ \text{J/K}} - 1\right) = 5.015.$$

Substituting this value for d into Eq. 20.26, I get $\gamma = 1 + \frac{2}{5.015} = 1.399$. ✔

(*b*) Because the nitrogen molecule is diatomic, it has three translational, two rotational, and two vibrational degrees of freedom. I know from Section 20.3 that vibrational motion does not begin to contribute to the heat capacity until the temperature of a gas is about 1000 K. Because my sample is at room temperature, and given that my result for d is very close to 5, I conclude that the translational and rotational degrees of freedom are the ones that contribute to the heat capacity ratio. ✔

❹ **EVALUATE RESULT** I can check the reasonableness of my calculated value $\gamma = 1.399$ by using it in Eq. 20.26 to determine C_V for the nitrogen molecule and then comparing what I get with the value obtained using data from Table 20.3. From Eq. 20.26: $C_V = C_P/\gamma = (4.844 \times 10^{-23}\ \text{J/K})/1.40 = 3.46 \times 10^{-23}$ J/K; from Table 20.3: $C_V = 2.50k_B = 2.50(1.381 \times 10^{-23}\ \text{J/K}) = 3.45 \times 10^{-23}$ J/K. That these two calculations result in about the same value for C_V gives me confidence in the value I obtained for γ.

The nitrogen molecule, N_2, is similar to the hydrogen molecule, H_2. I therefore expect the contribution of the various degrees of freedom to the heat capacity of the two molecules to be similar. I know from Section 20.3 that it is translational and rotational motions that contribute in H_2, which agrees with the motion types I chose in part *b*.



20.17 Suppose the isobaric process in Figure 20.34*c* begins and ends on the same isotherms as the isochoric process in Figure 20.30*c*. Which process requires a greater amount Q of energy transferred thermally?

An example of an isothermal (constant temperature) quasistatic process is shown in **Figure 20.35*a*** on the next page. Sand is dropped onto a piston, which then moves downward, compressing an ideal gas sample. The gas sample consists of N particles and is in thermal contact with a large volume of water. To keep the gas at constant temperature, the water volume must be large enough that adding (or removing) energy thermally to it does not noticeably change the water temperature. Such a body of water is called a thermal reservoir. The thermal reservoir is not part of the system (the gas by itself). The purpose of the thermal reservoir is to keep the temperature of our system constant by thermally exchanging energy with it.

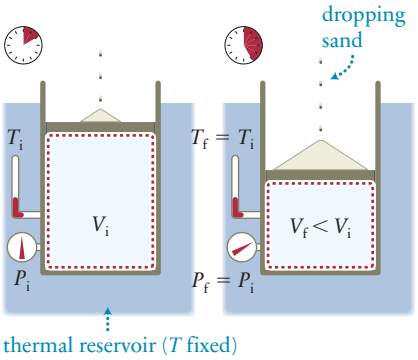
The work done by the piston on the gas is converted to thermal energy. For the process to be isothermal, the work has to be done sufficiently slowly for the gas and the thermal reservoir to remain in thermal equilibrium. A quasistatic isothermal process is thus a slow interplay between work done on the gas and thermal transfer of energy out of the gas.

The temperature of the gas remains constant, and hence the thermal energy does not change:

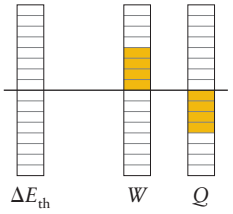
$$\Delta E_{\text{th}} = 0 \quad \text{(isothermal process)}. \tag{20.27}$$

Figure 20.35 An example of an isothermal quasistatic process.

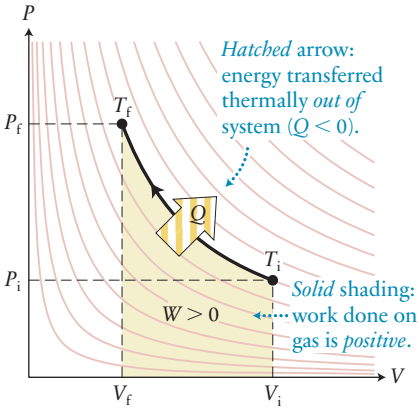
(*a*) Ideal gas compressed slowly at constant temperature



(*b*) Energy diagram for gas



(*c*) *PV* diagram showing the isothermal process



The energy diagram for the volume of ideal gas during this isothermal process is shown in Figure 20.35*b*. The piston is doing work on the gas ($W > 0$), and energy is transferred thermally from the gas ($Q < 0$). Because the temperature is constant, there is no change in the thermal energy. Hence the energy law for an ideal gas is $0 = W + Q$, which means

$$Q = -W \quad \text{(isothermal process)}. \tag{20.28}$$

The work done on the gas can be found by substituting $P = Nk_BT/V$ from the ideal gas law into Eq. 20.8:

$$W = -\int_{V_i}^{V_f} \frac{Nk_BT}{V}\,dV = -Nk_BT\int_{V_i}^{V_f} \frac{dV}{V} \quad \text{(isothermal process)}, \tag{20.29}$$

where we have pulled the (constant) temperature T out of the integral. Solving the integral, we have

$$W = -Nk_B\mathrm{T}(\ln V_f - \ln V_i) = -Nk_BT\ln\left(\frac{V_f}{V_i}\right) \quad \text{(isothermal process)}. \tag{20.30}$$

The energy transferred thermally to the gas is then, from Eq. 20.28,

$$Q = -W = Nk_BT\ln\left(\frac{V_f}{V_i}\right) \quad \text{(isothermal process)}. \tag{20.31}$$

Figure 20.35*c* shows the *PV* diagram for the isothermal process. The path lies along an isotherm. The shading for the arrow representing Q is hatched because energy is transferred thermally out of the gas. Note that $V_f < V_i$, and so $Q < 0$, as we concluded earlier (see Checkpoint 20.11).

Table 20.6 summarizes the expressions we have obtained for the various terms in the energy law for an ideal gas for each of the four constrained ideal gas processes.

Table 20.6 Energy law for an ideal gas and constrained ideal gas processes

Process	Constraint	W	Q	ΔE_{th}	Energy law for ideal gas
isochoric	V fixed	0	$NC_V\Delta T$		$\Delta E_{th} = Q$
isentropic	S fixed		0	$NC_V\Delta T$	$W = \Delta E_{th}$
isobaric	P fixed	$-Nk_B\Delta T$	$NC_P\Delta T$	$NC_V\Delta T$	$\Delta E_{th} = W + Q$
isothermal	T fixed	$-Nk_BT\ln\left(\frac{V_f}{V_i}\right)$		0	$Q = -W$

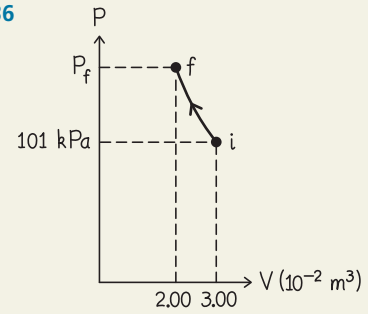
EXAMPLE 20.8 Isothermal compression

An ideal gas sample is compressed quasistatically at constant temperature. The initial pressure and volume are $P_i = 1.01 \times 10^5$ Pa and $V_i = 3.00 \times 10^{-2}$ m^3, and the final volume is $V_f = 2.00 \times 10^{-2}$ m^3. How much work is done on the gas during the compression?

❶ **GETTING STARTED** I begin by making a *PV* diagram of the process (**Figure 20.36**). The process is an isothermal compression, so it is represented by a path going up along an isotherm toward the left.

QUANTITATIVE TOOLS

Figure 20.36



2 DEVISE PLAN The work done on an ideal gas in an isothermal process is given by Eq. 20.30. I know the initial and final volumes, but not the factor Nk_BT. I can determine this factor from the ideal gas law, $PV = Nk_BT$; it does not matter whether I use initial values or final values in my calculation because the temperature is the same in both cases. Because I do not know the final pressure, however, I should use initial values.

3 EXECUTE PLAN From the ideal gas law, I have $Nk_BT = P_iV_i$. From Eq. 20.30, the work done on the gas is

$$W = -Nk_BT\ln\left(\frac{V_f}{V_i}\right) = -P_iV_i\ln\left(\frac{V_f}{V_i}\right).$$

Substituting values for the variables, I get

$$W = -(1.01\times10^5\text{ Pa})(3.00\times10^{-2}\text{ m}^3)\ln\left(\frac{2.00\times10^{-2}\text{ m}^3}{3.00\times10^{-2}\text{ m}^3}\right)$$
$$= 1.23\times10^3\text{ J}. \checkmark$$

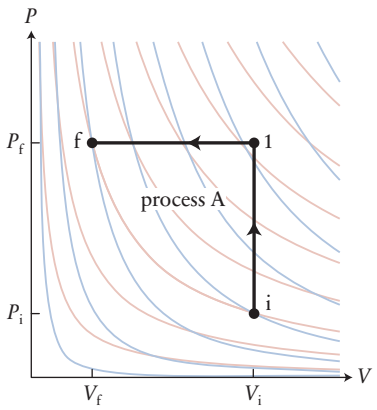
4 EVALUATE RESULT My answer is positive, as I expect for the work done on a gas in compressing it. To see whether the magnitude of my answer is in the right ballpark, I assume that the volume change is achieved by moving a piston of surface area 1×10^{-2} m^2 over a distance of 1 m. At a pressure of 1.01×10^5 Pa, the gas exerts on the piston a force of magnitude $(1\times10^{-2}\text{ m}^2)(1.01\times10^5\text{ Pa}) \approx 10^3$ N. The piston exerts a force of equal magnitude on the gas. If this force were constant, the work done on the gas during a force displacement of 1 m would be $W = (10^3\text{ N})(1\text{ m}) = 10^3$ J, which is of the same order of magnitude as the answer I obtained. (This estimate doesn't take into account the increase in pressure and so is a bit lower than the actual value.)



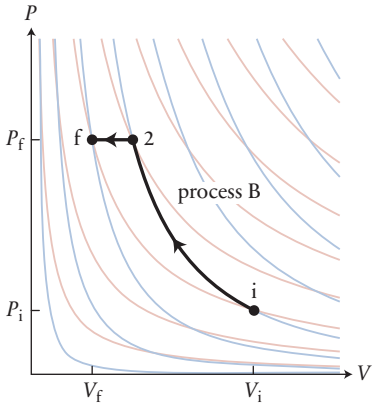
20.18 Consider an ideal gas undergoing the two processes depicted in Figure 20.37 between the same initial and final states that lie on an isotherm. In process A the gas is taken via an isochoric process to an intermediate state 1 and then via an isobaric process to the final state. In process B the gas undergoes an isentropic transformation to an intermediate state 2 and then an isobaric transformation to the final state. Which process has the greater value for (*a*) the work done on the gas and (*b*) the magnitude of the energy transferred thermally to the gas?

Figure 20.37 Checkpoint 20.18.

(*a*) Process A: isochoric step, then isobaric step



(*b*) Process B: isentropic step, then isobaric step



20.8 Entropy change in ideal gas processes

In this section we examine the entropy changes that occur for the four constrained quasistatic processes discussed in the preceding two sections. Recall from Chapter 19 that the entropy of a monatomic ideal gas is determined completely by the volume and the temperature of the gas (Eq. 19.59):

$$S = N\ln V + \tfrac{3}{2}N\ln T + \text{constant} \quad \text{(monatomic ideal gas)}. \qquad (20.32)$$

We obtained this expression for a monatomic ideal gas where each atom has three degrees of freedom (Section 20.3). The factor 3 in the $\frac{3}{2}$ in this expression comes from the three dimensions of the velocity space among which we equipartitioned the energy (Eq. 19.29). In the more general case of an ideal gas with d degrees of freedom, we replace the 3 by d:

$$S = N\ln V + \frac{d}{2}N\ln T + \text{constant} \quad \text{(ideal gas)}. \qquad (20.33)$$

During a process in which the gas undergoes a transformation from an initial equilibrium state (V_i, T_i) to a final equilibrium state (V_f, T_f), the entropy changes by an amount (see Eqs. 19.8 and 19.56)

$$\Delta S = N \ln\left(\frac{V_f}{V_i}\right) + \frac{d}{2} N \ln\left(\frac{T_f}{T_i}\right) \quad \text{(ideal gas)}. \tag{20.34}$$

For an isothermal process, $T_i = T_f$, and so the second term on the right is zero because $\ln(1) = 0$. Equation 20.34 thus becomes

$$\Delta S = N \ln\left(\frac{V_f}{V_i}\right) \quad \text{(isothermal process)}, \tag{20.35}$$

a result we already obtained in Eq. 19.8. Comparing this result with Eq. 20.31, we see that we can express the entropy change for an isothermal process as

$$\Delta S = \frac{Q}{k_B T} \quad \text{(isothermal process)}. \tag{20.36}$$

This expression allows us to calculate the change in entropy directly once we know the amount of energy transferred thermally, allowing us to bypass Eq. 20.34. Equation 20.36 is applicable only when the temperature is constant, however. For any other process, Eq. 20.34 provides the simplest approach for calculating the entropy change.

For an isochoric process, $V_f = V_i$, and so Eq. 20.34 reduces to

$$\Delta S = \frac{d}{2} N \ln\left(\frac{T_f}{T_i}\right) \quad \text{(isochoric process)} \tag{20.37}$$

or, substituting from Eq. 20.14,

$$\Delta S = \frac{N C_V}{k_B} \ln\left(\frac{T_f}{T_i}\right) \quad \text{(isochoric process)}. \tag{20.38}$$



20.19 Can you use Eq. 20.36 to calculate the change in entropy for an isochoric process?

In an isobaric process, both V and T change, but the two are related by the ideal gas law, $PV = Nk_BT$. Because P is constant in any isobaric process, we have $V_i = (Nk_B/P)T_i$ and $V_f = (Nk_B/P)T_f$. Therefore

$$\frac{V_f}{V_i} = \frac{(Nk_B/P)T_f}{(Nk_B/P)T_i} = \frac{T_f}{T_i} \quad \text{(isobaric process)}. \tag{20.39}$$

With this result, Eq. 20.34 becomes

$$\begin{aligned} \Delta S &= N \ln\left(\frac{T_f}{T_i}\right) + \frac{d}{2} N \ln\left(\frac{T_f}{T_i}\right) \\ &= N\left(1 + \frac{d}{2}\right) \ln\left(\frac{T_f}{T_i}\right) \quad \text{(isobaric process)}. \end{aligned} \tag{20.40}$$

Substituting from Eq. 20.25, we can rewrite this expression in the form

$$\Delta S = \frac{N C_P}{k_B} \ln\left(\frac{T_f}{T_i}\right) \quad \text{(isobaric process)}. \tag{20.41}$$

Note the similarity between the expression for the entropy change in an isochoric process (Eq. 20.38) and the expression for the entropy change in an isobaric process (Eq. 20.41). The only difference is the type of heat capacity per particle: For the isochoric process we use C_V; for the isobaric process we use C_P.

As an example, consider the isobaric expansion shown in Figure 20.34. The final temperature is higher than the initial temperature, and so, because $\ln(T_f/T_i) > 0$, Eq. 20.41 tells us that the entropy increases. Indeed, from the PV diagram in Figure 20.34*c* we see that the final state must lie on a higher isentrope than the initial state, indicating that the entropy of the final state is greater than the entropy of the initial state.

In an isentropic process, the entropy doesn't change, and so

$$\Delta S = 0 \quad \text{(isentropic process)}. \tag{20.42}$$

This result implies that during an isentropic process the changes in volume and temperature are such that the two terms in Eq. 20.34 cancel each other. In other words, the volume and temperature changes are correlated in such a way as to keep the entropy of the system constant.



20.20 Suppose two different combinations of constrained quasistatic processes take two identical ideal gases from the same initial equilibrium state to the same final equilibrium state. Is the change in entropy the same or different in the two cases?

Table 20.7 summarizes the entropy changes for the four constrained ideal gas processes. As you saw in Checkpoint 20.20, the value of the entropy change is independent of the process and depends on only the temperatures and volumes of the initial and final equilibrium states. For *any* process, we can use Eq. 20.34 to calculate the entropy change. The expressions summarized in Table 20.7, however, provide a quick way to obtain the entropy changes for the four basic constrained ideal gas processes. Keep in mind that in these processes the system is not closed, and so the entropy change can be positive, negative, or zero.

Equation 20.33 can be rewritten in a form that allows us to derive useful relationships among the pressure, volume, and temperature of any two states, (P_i, V_i, T_i) and (P_f, V_f, T_f), that lie on an isentrope. We begin by combining the first and second terms on the right in Eq. 20.33, which gives us

$$S = N \ln(T^{d/2}V) + \text{constant}. \tag{20.43}$$

Table 20.7 Entropy changes in ideal gas processes

Process	ΔS	Equation
isochoric	$\frac{NC_V}{k_B} \ln\left(\frac{T_f}{T_i}\right)$	20.38
isothermal	$N \ln\left(\frac{V_f}{V_i}\right) = \frac{Q}{k_B T}$	20.35, 20.36
isobaric	$\frac{NC_P}{k_B} \ln\left(\frac{T_f}{T_i}\right)$	20.41
isentropic	0	20.42

Because the entropy is constant, the quantity $T^{d/2}V$ must be constant as well for a sample that consists of a fixed number N of particles: $T_i^{d/2}V_i = T_f^{d/2}V_f$. Raising the constant quantity $T^{d/2}V$ to the power $\frac{2}{d}$, and substituting $\frac{2}{d} = \gamma - 1$ from Eq. 20.26, we see that $(T^{d/2}V)^{2/d} = TV^{\gamma-1}$, or

$$T_iV_i^{\gamma-1} = T_fV_f^{\gamma-1} \quad \text{(isentropic process)}. \tag{20.44}$$

This expression provides the relationship between volume and temperature along an isentrope. Using the ideal gas law, we can obtain expressions relating other quantities along an isentrope. Multiplying the ideal gas law, $PV = Nk_BT$, by $V^{\gamma-1}$, we obtain:

$$PV^{\gamma} = Nk_BTV^{\gamma-1}. \tag{20.45}$$

From Eq. 20.44 we see that the right side of Eq. 20.45 is constant for an isentropic process, and thus the quantity PV^{γ} must also be constant:

$$P_iV_i^{\gamma} = P_fV_f^{\gamma} \quad \text{(isentropic process)}. \tag{20.46}$$

From the ideal gas law, we know that V is proportional to T/P, which means that $PV^{\gamma} \propto P(T/P)^{\gamma} = P^{1-\gamma}T^{\gamma}$ is constant for an isentropic process. Raising this quantity to the power $1/\gamma$, we get $(P^{1-\gamma}T^{\gamma})^{1/\gamma} = P^{(1/\gamma)-1}T$, and so in an isentropic process,

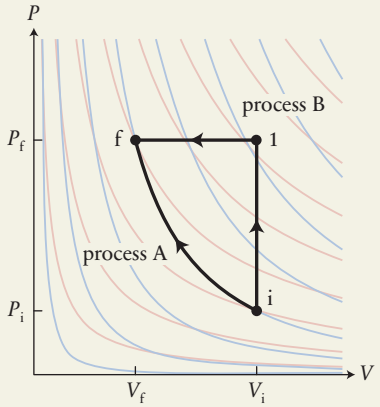
$$P_i^{(1/\gamma)-1}T_i = P_f^{(1/\gamma)-1}T_f \quad \text{(isentropic process)}. \tag{20.47}$$

Equations 20.44, 20.46, and 20.47 relate the pressure, volume, and temperature for the initial and final equilibrium states of an ideal gas sample that consists of N particles undergoing an isentropic process.

EXAMPLE 20.9 Entropy changes

Figure 20.38 shows two quasistatic processes that take a sample of an ideal gas containing N particles from an initial state (P_i, V_i) to a final state (P_f, V_f). Process A is a one-step isentropic compression. Process B has two steps: an isochoric pressure increase to an intermediate state 1 followed by an isobaric decrease in volume. Express the changes in entropy for process A and process B in terms of the given quantities.

Figure 20.38



❶ **GETTING STARTED** The fact that the initial and final states are on different isotherms (red curves) in Figure 20.38 tells me that the two are at different temperatures, which I call T_i and T_f. These temperatures are related to the initial and final volume and pressure by the ideal gas law. For the intermediate state, I can write $(P_1, V_1) = (P_f, V_i)$. I'll denote the temperature of this state by T_1.

❷ **DEVISE PLAN** Table 20.7 gives expressions for the entropy changes in the four constrained processes. I can eliminate the temperature from these expressions by using the ideal gas law, $PV = Nk_BT$, and Eq. 20.39.

❸ **EXECUTE PLAN** Process A is isentropic, which means that its entropy doesn't change: $\Delta S_A = 0$. ✔

I first obtain the entropy change for the isochoric step of process B from Eq. 20.38:

$$\Delta S_{B,\text{isochoric}} = \frac{NC_V}{k_B}\ln\left(\frac{T_1}{T_i}\right).$$

From the ideal gas law, I get for the initial and intermediate states, respectively:

$$P_iV_i = Nk_BT_i \tag{1}$$

$$P_fV_i = Nk_BT_1. \tag{2}$$

Dividing Eq. 2 by Eq. 1 gives me $T_1/T_i = P_f/P_i$, and so for the first step of process B I have

$$\Delta S_{B,\text{isochoric}} = \frac{NC_V}{k_B}\ln\left(\frac{P_f}{P_i}\right). \tag{3}$$

The entropy change for the isobaric step of process B is given by Eq. 20.41:

$$\Delta S_{B,\text{isobaric}} = \frac{NC_P}{k_B}\ln\left(\frac{T_f}{T_1}\right).$$

Equation 20.39 yields $T_f/T_1 = V_f/V_i$, so I can write the entropy change as

$$\Delta S_{B,\text{isobaric}} = \frac{NC_P}{k_B}\ln\left(\frac{V_f}{V_i}\right). \tag{4}$$

The entropy change for process B is the sum of the entropy changes for the two steps, so, adding Eqs. 3 and 4, I get

$$\Delta S_B = \frac{NC_V}{k_B}\ln\left(\frac{P_f}{P_i}\right) + \frac{NC_P}{k_B}\ln\left(\frac{V_f}{V_i}\right). \checkmark$$

4 EVALUATE RESULT The two processes connect the same initial and final states, so the entropy change should be the same for the two processes. Therefore the expression I obtained for process B should be zero: $\Delta S_B = 0$. Because the process that takes the gas from the initial to the final state is isentropic, Eq. 20.46 tells me that $P_f/P_i = (V_i/V_f)^\gamma$. Summing my expressions for $\Delta S_{B,\text{isochoric}}$ and $\Delta S_{B,\text{isobaric}}$ and making this substitution for P_f/P_i yield

$$\Delta S_B = \frac{NC_V}{k_B}\ln\left(\frac{V_i}{V_f}\right)^\gamma + \frac{NC_P}{k_B}\ln\left(\frac{V_f}{V_i}\right)$$

$$= \frac{NC_V}{k_B}\gamma\ln\left(\frac{V_i}{V_f}\right) + \frac{NC_P}{k_B}\ln\left(\frac{V_f}{V_i}\right).$$

Because $\gamma \equiv C_P/C_V$ (Eq. 20.26), I obtain, as expected

$$\Delta S_B = \frac{NC_P}{k_B}\ln\left(\frac{V_i}{V_f}\right) + \frac{NC_P}{k_B}\ln\left(\frac{V_f}{V_i}\right)$$

$$= \frac{NC_P}{k_B}\ln\left(\frac{V_i}{V_f}\right) - \frac{NC_P}{k_B}\ln\left(\frac{V_i}{V_f}\right) = 0.$$



20.21 For a quasistatic ideal gas process that ends on a higher isentrope, are W, Q, ΔE_{th}, and ΔS positive, negative, or zero when the gas is (*a*) expanded and (*b*) compressed?

20.9 Entropy change in nonideal gas processes

Let us now briefly consider what happens to the energy and entropy when the temperature of a liquid or gas changes. When we deal with a solid or liquid of given mass and constant volume, it is more convenient to work with specific heat capacity (see Section 20.3 and Table 20.2) rather than heat capacity per particle. The specific heat capacity is obtained by dividing the energy transferred thermally by the mass of the object or system and by the temperature change:*

$$c_V \equiv \frac{Q}{m\Delta T}. \tag{20.48}$$

And so for a given change in temperature, the energy transferred thermally is

$$Q = mc_V\Delta T. \tag{20.49}$$

Because the volume is constant, no work is done and the energy law reduces to $\Delta E_{th} = Q$, so the system's thermal energy change is

$$\Delta E_{th} = mc_V\Delta T. \tag{20.50}$$

*We are assuming here that the specific heat capacity does not depend on temperature. As Figure 20.15 shows, however, this is only approximately true over limited temperature ranges.

To calculate the corresponding entropy change, we rewrite the definition of absolute temperature (Eq. 19.38), which gives the relationship between changes in thermal energy and entropy, as follows:

$$dS = \frac{dE_{\text{th}}}{k_{\text{B}}T}. \tag{20.51}$$

For an infinitesimally small change in temperature, Eq. 20.50 becomes $dE_{\text{th}} = mc_V dT$, and so Eq. 20.51 becomes

$$dS = \frac{mc_V}{k_{\text{B}}}\frac{dT}{T}. \tag{20.52}$$

To determine the entropy change for a finite temperature change, we integrate this expression for a temperature change $\Delta T = T_{\text{f}} - T_{\text{i}}$:

$$\Delta S = \frac{mc_V}{k_{\text{B}}}\int_{T_{\text{i}}}^{T_{\text{f}}}\frac{dT}{T} = \frac{mc_V}{k_{\text{B}}}\ln\left(\frac{T_{\text{f}}}{T_{\text{i}}}\right). \tag{20.53}$$

We can also use Eq. 20.51 to obtain a more general expression for the entropy change in an isothermal process. Let the energy thermally transferred be Q. At constant temperature the entropy change is then

$$\Delta S = \frac{Q}{k_{\text{B}}T} \quad \text{(isothermal process)}. \tag{20.54}$$

Note that this expression is identical to the one we obtained earlier for an ideal gas (Eq. 20.36).

Let us next determine what happens to the energy and entropy of a system during a phase transition. In Section 20.3 we defined the **specific transformation energy** as the energy transferred thermally to a substance undergoing a phase transition divided by the mass of the substance:

$$L \equiv \frac{Q}{m} \quad \text{(phase transition with } Q > 0\text{)}. \tag{20.55}$$

Because the quantity L is always positive and Q can be negative or positive, Eq. 20.55 holds only for melting and vaporization. For solidification and condensation, Q is negative and so we have

$$L \equiv \frac{-Q}{m} \quad \text{(phase transition with } Q < 0\text{)}. \tag{20.56}$$

Because the temperature is constant, we can use Eq. 20.54 to calculate the corresponding entropy change

$$\Delta S = \frac{Q}{k_{\text{B}}T} = \frac{mL}{k_{\text{B}}T} \quad \text{(melting or vaporization)}. \tag{20.57}$$

For solidification and condensation the entropy change is of the opposite sign:

$$\Delta S = -\frac{mL}{k_{\text{B}}T} \quad \text{(solidification or condensation)}. \tag{20.58}$$

QUANTITATIVE TOOLS



20.22 An amount of water with a mass of 0.010 kg at a temperature of 20 °C is placed in a freezer and turns into an ice cube at 0.0 °C. What is the corresponding entropy change? Ignore any volume changes.

Chapter Summary

Thermal interactions and temperature (Sections 20.1, 20.2, 20.5)

Concepts During a **thermal interaction,** thermal energy is transferred from an object at a higher temperature to an object at a lower temperature. The energy Q transferred is called **energy transferred thermally.**

During an **adiabatic process,** only mechanical energy is transferred (work is done on the system); no energy is transferred thermally, so $Q = 0$.

When a gas volume decreases, positive work is done on the gas ($W > 0$); when the volume increases, negative work is done on the gas ($W < 0$).

During a **quasistatic process,** the system always remains close to equilibrium. If a quasistatic process is carried out adiabatically ($Q = 0$), the entropy of the system does not change ($\Delta S = 0$).

At the **triple point of water,** the solid, liquid, and gas phases coexist in thermal equilibrium. This point occurs at 273.16 K and 610 Pa.

Quantitative Tools When the kinetic, potential, and source energies of an ideal gas are constant, only a change in the thermal energy can change the energy of the gas, and this change is given by

$$\Delta E_{\text{th}} = W + Q. \tag{20.2}$$

The work done on a gas during a quasistatic process is

$$W = -\int_{V_i}^{V_f} P\, dV. \tag{20.8}$$

Conversions between the Celsius, Kelvin, and Fahrenheit temperature scales:

$$T_C = \frac{1\ ^\circ\text{C}}{1\ \text{K}}\, T - 273.15\ ^\circ\text{C}$$

$$T_C = \frac{5\ ^\circ\text{C}}{9\ ^\circ\text{F}}(T_F - 32\ ^\circ\text{F}).$$

Thermodynamic processes (Sections 20.3–20.7)

Concepts The **heat capacity** (J/K) of an *object* is a scalar defined as the amount of energy transferred thermally to an object divided by the resulting change in temperature of the object.

The **specific heat capacity** (J/K · kg) is a scalar equal to the heat capacity of an object divided by the mass of the object.

A **degree of freedom** is any independent way of storing energy in a particle.

In an ideal gas in thermal equilibrium, each particle's **equipartition energy share** of the energy of the gas is $dk_BT/2$, where d is the number of degrees of freedom available to the particles.

A PV diagram is a graph of the pressure in a gas as a function of the gas volume. Any point on a PV diagram represents an equilibrium state in which the gas has a well-defined temperature, thermal energy, and entropy. Any continuous path on a PV diagram represents a quasistatic process.

A **thermal reservoir** is a body of material large enough that its temperature does not change in a thermal interaction.

An **isothermal process** is one that occurs at constant temperature. An **isobaric process** is one that occurs at constant pressure. An **isochoric process** is one that occurs at constant volume. An **isentropic process** is one in which the entropy does not change.

Quantitative Tools At temperature T, the thermal energy of a sample of an ideal gas that contains N particles, each having d degrees of freedom, is

$$E_{\text{th}} = \frac{d}{2} N k_B T, \tag{20.4}$$

where $k_B = 1.381 \times 10^{-23}$ J/K is the Boltzmann constant.

At constant volume, the **heat capacity per particle** of a gas that contains N particles, each having d degrees of freedom, is

$$C_V \equiv \frac{Q}{N\Delta T} = \frac{d}{2} k_B. \tag{20.13) (20.14}$$

At constant pressure, the **heat capacity per particle** is

$$C_P \equiv \frac{Q}{N\Delta T} = C_V + k_B = \left(\frac{d}{2} + 1\right) k_B. \tag{20.20) (20.24) (20.25}$$

For any ideal gas process,

$$\Delta E_{\text{th}} = N C_V \Delta T. \tag{20.15}$$

The unitless **heat capacity ratio** γ is

$$\gamma \equiv \frac{C_P}{C_V} = 1 + \frac{2}{d}. \tag{20.26}$$

Table 20.4 Phase transition temperatures and specific transformation energies for melting (m) and vaporization (v)

Substance	T_m (K)	L_m (10^3 J/kg)	T_v (K)	L_v (10^3 J/kg)
Helium	–	–	4.230	20.8
Hydrogen	13.84	58.6	20.37	449
Nitrogen	63.18	25.5	77.36	199
Oxygen	54.36	13.8	90.19	213
Ammonia	195	331	239.8	1369
Carbon tetrachloride	250.3	16	349.8	194
Ethanol	159	100.7	351	837
Water	273.15	334	373.12	2256
Mercury	234	11.8	630	272
Aluminum	933.6	395	2740	10,500
Lead	600.5	24.5	2023	871
Sulfur	392	38.1	717.75	326
Gold	1336.15	64.5	2933	1578
Copper	1356	134	2840	5069

Entropy in ideal gas processes (Sections 20.2, 20.8)

Concepts At **absolute zero** ($T = 0$), the thermal energy of a system is a minimum and its entropy is zero because it has only one basic state available to it.

Quantitative Tools The entropy of an ideal gas that has d degrees of freedom is

$$S = N \ln V + \frac{d}{2} N \ln T + \text{constant}, \tag{20.33}$$

and the change in its entropy during a process is

$$\Delta S = N \ln\left(\frac{V_f}{V_i}\right) + \frac{d}{2} N \ln\left(\frac{T_f}{T_i}\right). \tag{20.34}$$

Table 20.6 Energy law for an ideal gas and constrained ideal gas processes

Process	Constraint	W	Q	ΔE_{th}	Energy law for ideal gas
isochoric	V fixed	0	$NC_V\Delta T$		$\Delta E_{th} = Q$
isentropic	S fixed		0	$NC_V\Delta T$	$W = \Delta E_{th}$
isobaric	P fixed	$-Nk_B\Delta T$	$NC_P\Delta T$	$NC_V\Delta T$	$\Delta E_{th} = W + Q$
isothermal	T fixed	$-Nk_BT \ln\left(\frac{V_f}{V_i}\right)$		0	$Q = -W$

For any isentropic process:

$$T_i V_i^{\gamma-1} = T_f V_f^{\gamma-1} \tag{20.44}$$

$$P_i V_i^{\gamma} = P_f V_f^{\gamma} \tag{20.46}$$

$$P_i^{\frac{1}{\gamma}-1} T_i = P_f^{\frac{1}{\gamma}-1} T_f. \tag{20.47}$$

Entropy in nonideal gas processes (Section 20.9)

Quantitative Tools For a solid or liquid that has a specific heat capacity c_V, the entropy change when a sample of mass m undergoes a temperature change is

$$\Delta S = \frac{mc_V}{k_B} \ln\left(\frac{T_f}{T_i}\right). \quad (20.53)$$

The entropy change during a melting or vaporization phase transition at temperature T is

$$\Delta S = \frac{mL}{k_B T}. \quad (20.57)$$

The entropy change during a solidification or condensation phase transition at temperature T is

$$\Delta S = -\frac{mL}{k_B T}. \quad (20.58)$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

Unless you are told otherwise, assume that all processes in these problems are quasistatic and that all gases are monatomic and ideal.

20.1 Thermal interactions

1. When the volume of a gas increases, is the work done by the gas on its surroundings positive, negative, or zero? •
2. A thermally insulated chamber fitted with a piston as its bottom surface contains an ideal gas. The piston slowly moves upward and compresses the gas into a smaller volume. Draw an energy diagram for this process, taking the gas as your system. •
3. State whether each process is quasistatic, adiabatic, both, or neither for the defined system: (*a*) Water boils in a pot. System: water. (*b*) A balloon filled with air is slowly compressed and cooled. System: balloon and air. (*c*) An initially sealed vial of air is opened in an air-filled room that is isolated from its outside environment, where the temperature and number density of air particles are the same inside and outside the vial. System: vial, room, and air. •
4. A sample that contains 5.60×10^{21} particles of a monatomic ideal gas expands adiabatically and quasistatically from 1.00 L to 2.00 L. If the initial gas pressure is 30.5 kPa, what is the final temperature of the gas? ••
5. You are designing an engine that contains several pistons that each move 150 mm per stroke. You believe the engine may consume less energy than a traditional engine if many of the processes in the engine are reversible. The gas is a mixture of several substances, but the particles have an average mass of 7.548×10^{-26} kg per particle, and the gas temperature can reach 300 °F. Determine an upper limit on the number of cycles a piston can go through each second, and satisfy the reversibility constraint. How could this upper limit be increased? Is the reversibility constraint sensible? •••

20.2 Temperature measurement

6. A Kelvin thermometer and a Fahrenheit thermometer both give the same reading for a certain sample at what numerical value? •
7. You poured freshly brewed hot tea into a cup. A thermometer indicates that its temperature is 96.8 °C. Convert this temperature into the Fahrenheit and Kelvin scales. •
8. You are monitoring the temperature of a solution in the laboratory. Using a thermometer, you recorded a temperature of 40.5 °C. Two hours later, about to do the same measurement, you accidentally dropped the thermometer you previously used and broke it. Instead, you used a differently scaled thermometer for the measurement and recorded a reading of 144 °F. Express the difference between the two temperature readings in (*a*) degrees Celsius, (*b*) kelvin, and (*c*) degrees Fahrenheit. ••
9. While converting a list of degrees Celsius temperatures to degree Fahrenheit, you notice that both scales have the same numerical value at a certain reading. What is this number? •
10. When a certain ideal gas thermometer is placed in water at the freezing point, the mercury level in the right arm is 869 mm above the reference mark. How far is the mercury level above the reference mark when this thermometer is placed at room temperature where the atmospheric pressure is 1.00 atm (See Table 20.1 for temperature values.)? ••
11. When a certain ideal glass thermometer is at room with a temperature of 25 °C, the mercury level in the right arm is 245 mm above the reference mark. When the thermometer is then dipped into a saline solution, the mercury level drops by 25 mm. (*a*) What is the temperature of saline solution? (*b*) Is your answer to part *a* sufficient enough to solidify a certain type of oil with a freezing point of 2.77 °C? ••
12. When a certain ideal gas thermometer is immersed in water at the triple point, the mercury level in the right arm is 124 mm above the reference mark. When this thermometer is then immersed in a liquid of unknown temperature, the mercury level drops by 68.0%. (*a*) What is

the temperature of the liquid? (*b*) What is a potential source of error in this measurement? ●●

13. An engineer planning a weld from manufacturer specifications realizes that the units for temperature have been accidentally omitted. The specifications say, "The temperature of the weld should be no lower than 95% of 1960 but can be safely welded at a much higher temperature." The engineer is not sure whether the 1960 is 1960 K or 1960 °C. To what temperature (in degrees Celsius) should she specify the weld to be sure it is not done at too low a temperature? ●●
14. For an environmental study, you need to collect temperature data for a year in a North American desert, where the temperature ranges from 255 K in winter to 320 K in summer. You have a single thermal sensor that can be calibrated in degrees Celsius or degrees Fahrenheit, but the readout is fixed to one decimal place. (*a*) Which scale, Fahrenheit or Celsius, will give you greater precision? (*b*) Your research group doubts your choice, so you must provide numerical evidence to support it. What is the ratio of the number of readings possible on the scale you chose to the number of readings possible on the scale you didn't choose? ●●●

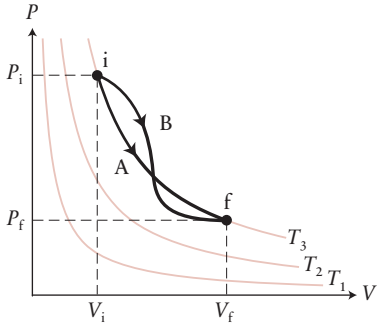
20.3 Heat capacity

15. How much energy is needed to boil 750 g of water at freezing point? ●
16. A 250-g piece of an unknown metal has been in boiling water for several minutes. To bring this metal at room temperature (20 °C), 89.8 kJ of heat must be removed. (*a*) Calculate the specific heat of the metal. (*b*) Determine what type of metal is the sample using your answer in part *a*. ●
17. According to the Can Manufacturers Institute, the energy used to make an aluminum can from recycled aluminum is 5% of the energy used to make an aluminum can from virgin ore. In a typical year, 1.7 billion pounds of aluminum cans are recycled. How much energy is thermally transferred to get this mass of aluminum from 20 °C to its melting point, 660 °C? ●●
18. Three diatomic molecules from different gas samples—one at 3 K, one at 298 K, and one at 1000 K—are in thermal equilibrium with their respective surroundings. (*a*) Rank the molecules according to heat capacity per molecule at constant volume, highest first. (*b*) What is the thermal energy of each molecule, expressed in terms of T and k_B? ●●
19. How much heat energy is needed to convert a mixture of 20.0 g of ice and 35.0 g of water at thermal equilibrium to steam at 100.0 °C? If this energy were used to lift a truck, to what height could a 1.00×10^4-kg truck be lifted? ●●
20. A 0.350-kg aluminum calorimeter contains 0.110 kg of water, and both are at room temperature (20.0°C). A 0.280-kg block of lead at 90.0°C is then dropped into the container. Assuming no heat loss to the surroundings, find the final temperature of the system. ●●
21. Devil's Throat, the longest drop in Iguazu Falls on the Brazil-Argentina border, has a height of 82 m. If all the kinetic energy a given volume of water acquires in dropping through Devil's Throat could go into heating the water, how much would the water temperature rise? ●●
22. You pour 5.0 L of water at 20 °C into a large pot to make spaghetti. If the heating element on your stove is rated at 1250 W, how many minutes does it take for the water to start boiling? Ignore any energy thermally transferred to the environment. ●●
23. A helium atom is in a nanowire (a wire with a diameter on the scale of molecules) in thermal equilibrium at 77 K. (*a*) The atom is constrained in the nanowire and so can move in only one dimension, along the wire length. What is the atom's thermal energy? (*b*) The atom is boiled off the nanowire and allowed to float free at 295 K. What is its thermal energy now? ●●●
24. A diatomic molecule in thermal equilibrium at 120 K is constrained to move freely only on a two-dimensional surface. (*a*) What is the molecule's thermal energy? (*b*) The molecule is boiled off the surface and allowed to float freely in a gas at thermal equilibrium at 298 K. What is the molecule's thermal energy now? (*c*) What is its thermal energy if it is heated to 3000 K? ●●
25. A cryogenic substance is found to have a specific heat capacity c that varies with temperature according to $c = \beta T^2$, where β is an empirically derived constant with units $\text{J/K}^3 \cdot \text{kg}$. If 231 J of energy must be transferred thermally to an 8000-mg sample of this substance to raise the temperature of the sample from 1.00 to 6.00 K, what is the value of β? ●●●

20.4 *PV* diagrams and processes

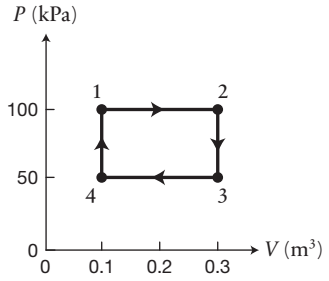
26. An ideal gas expands adiabatically and quasistatically in a chamber fitted with a piston. Is this process isothermal, isentropic, isobaric, isochoric, or none of these? ●
27. Figure P20.27 shows two processes, A and B, carried out on an ideal gas. In which process was the amount of work done by the gas greater? ●

Figure P20.27



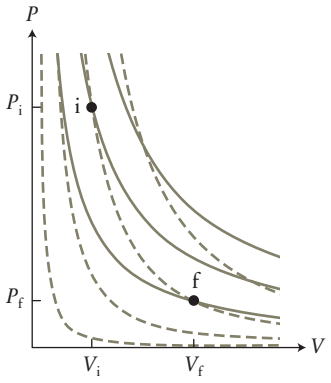
28. Figure P20.28 is the *PV* diagram for a process carried out on 10 mol of an ideal gas in thermal equilibrium that begins and ends on the same state (the initial state i and the final state f for the gas are at 1). (*a*) Characterize each leg of the process—$1 \rightarrow 2$, $2 \rightarrow 3$, $3 \rightarrow 4$, $4 \rightarrow 1$—as being isothermal, isobaric, isochoric, or isentropic. (*b*) Determine the temperature at state 1 and at state 3. ●●

Figure P20.28



29. The curves in Figure P20.29 are isochors, isobars, isotherms, or isentropes. If a system is taken from initial state i to final state f, with $V_i = 0.15\ m^3$ and $P_f = 1.00$ atm, what is the change in the entropy of the system? ●●

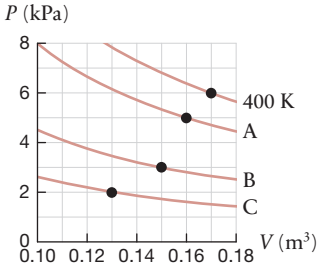
Figure P20.29



30. A 0.300-kg sample of nitrogen gas (diatomic molecules, $m_{N_2} = 4.652 \times 10^{-26}$ kg) in a chamber fitted with a piston undergoes an isothermal expansion from 0.0500 m^3 to 0.150 m^3. If the final pressure is 150 kPa, what is the final temperature? ●●

31. Figure P20.31 shows four isotherms on the *PV* diagram for an ideal gas. What are the temperatures of isotherms A, B, and C? ●●

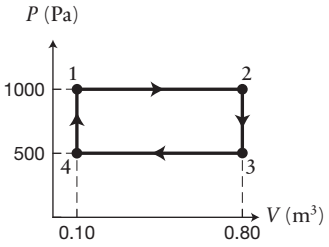
Figure P20.31



20.5 Change in energy and work

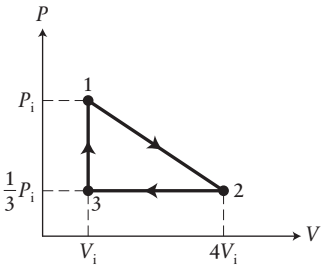
32. An ideal gas undergoes the process represented in Figure P20.32, where the gas is in state 1 when the process begins and again in state 1 at the finish. (*a*) Determine the work done by the gas during each of the four legs—$1 \rightarrow 2$, $2 \rightarrow 3$, $3 \rightarrow 4$, $4 \rightarrow 1$. (*b*) Determine the work done by the gas for the process. ●

Figure P20.32



33. An ideal gas undergoes the process represented by the *PV* diagram of Figure P20.33, where the gas is in state 1 when the process begins and again in state 1 at the end of the process. Calculate the work done on the gas as a function of the initial pressure and volume. ●

Figure P20.33



34. If 1.00 mol of an ideal monatomic gas initially at 72 K absorbs 100 J of thermal energy, what is the final temperature? Assume the volume does not change. ●

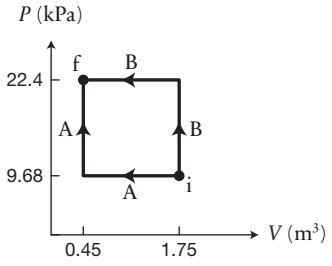
35. An ideal gas is inside a cylinder fitted with a piston of cross-sectional area 0.10 m^2, and initially the gas pressure is 5.0×10^4 Pa. A quantity $Q = 5.0$ kJ of thermal energy is transferred slowly to the gas while the piston moves 0.10 m and the gas expands. If the expansion is isobaric, how much work is done on the gas? ●●

36. What is the thermal energy associated with 5.6×10^{18} nitrogen molecules, N_2, at 100 °F? Assume the nitrogen can be treated as an ideal gas. ●●

37. In a cylindrical chamber fitted with a piston, the piston compresses 10.0 mol of ideal gas from 1.00 m^3 to 0.100 m^3 at a constant temperature of 300 K. Suppose this compression was done by placing a 100-kg block on top of the piston. What is the piston radius? Assume the air pressure outside the piston is 1.00 atm, and ignore the piston mass. ●●

38. Figure P20.38 shows two processes, A and B, for moving 3.45×10^{22} particles of a monatomic ideal gas from state i to state f. (*a*) Which process requires less work done on the gas? (*b*) Which process requires a smaller quantity Q of energy transferred thermally to the gas? (*c*) By how much does the work done on the gas in process A differ from the work done on the gas in process B? ••

Figure P20.38

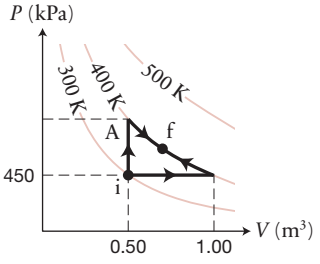


39. A 1.00-mol sample of an ideal diatomic gas in which the gas particles both translate and rotate is initially at 600 K. Energy is then added thermally to the sample until its temperature is 1000 K. At this temperature, the particles also vibrate. The sample is then heated to 1200 K. How much thermal energy does the sample absorb? ••
40. You have a sample that contains 9.70×10^{21} particles of an unknown gas in a chamber that has a volume of 0.0100 m^3. The initial pressure in the chamber is 1.84×10^4 Pa, and you determine that 378 J of energy must be added to the sample to increase the pressure to 2.68×10^4 Pa. (*a*) What can you say about the structure of the particles? (*b*) If your choices for the identity of the gas are helium, hydrogen, ammonia, and nitrogen, which do you choose, and why? •••

20.6 Isochoric and isentropic ideal gas processes

41. The PV diagram in Figure P20.41 shows two processes for taking an ideal gas from an initial state i to a final state f. How much work is done on the gas in the isochoric leg labeled A? •

Figure P20.41



42. You park your car in the sun with the windows rolled up, and the interior temperature rises from 35 °C at 100 kPa to 48 °C. The interior volume of your car is roughly 1.5 m × 2.0 m × 1.0 m. (*a*) What is the mass of the air in the car, assuming a mass density of 1.13 kg/m^3 for air at 35 °C? (*b*) How much energy was delivered from the Sun to the interior to raise the temperature as indicated if all the energy went into the air? Use 720 J/K · kg for the specific heat capacity of air. ••
43. Your one-story house has a floor area of 2.0×10^2 m^2 and 3.0-m-high ceilings. (*a*) On a hot day, you leave home with all the windows closed and the air-conditioning on. A power failure causes the air-conditioning to stop, and by the time you return home, the quantity of thermal energy you calculated in Problem 42 has been added to the inside air. By how many kelvins has the inside air temperature increased? (*b*) How much thermal energy from the Sun would be required to raise the temperature in your house from 24 °C to 25 °C? ••
44. Steam at 120 °C and 101.3 kPa is added to a hollow steel can that is then sealed and allowed to cool. (*a*) What is the pressure inside the can at 101 °C? (*b*) The can is 0.20 m tall and 0.10 m in diameter. What is the magnitude of the force exerted by the atmosphere on the curved and flat portions of the can exterior? (*c*) Once the steam cools sufficiently, it begins to condense into liquid water and the can collapses inward. Why does the can retain its shape for much, but not all, of this process? (Hint: Search "collapsing soda can" on your favorite Internet video site.)
45. A sample of helium gas is allowed to expand in a process that is adiabatic and quasistatic. As the gas cools from 105 °C to 101 °C, it does 9.05 J of work on a piston. How many helium atoms are there in the sample? ••

20.7 Isobaric and isothermal ideal gas processes

46. A 2.00-L sample of an ideal gas initially at 1.00 atm and 273 K undergoes an isobaric process that cools the sample to 265 K. (*a*) What is the final pressure in the gas? (*b*) What is the final volume of the gas? •
47. A cylinder fitted with a piston contains 1.00 mol of air at room temperature, 20.0 °C. The air is slowly compressed by the piston while the temperature is kept constant. What is the work done on the air if the final volume is one-third the initial volume? •
48. A chamber fitted with a piston can be controlled to keep the pressure in the chamber constant as the piston moves up and down to increase or decrease the chamber volume. If the chamber contains an ideal gas at 296 K and 1.00 atm, what is the work done on the gas as the piston compresses it from 1.00 L to 0.333 L? •
49. Your lab mate is measuring the heat capacity ratio of an unknown diatomic gas at room temperature and standard atmospheric pressure. She obtains values for C_P and C_V and tells you the ratio is 6/9. Give the reason you need to check your lab mate's work. ••
50. A spherical air bubble in a lake expands as it rises slowly to the surface. At the point it starts to rise, the pressure is 2.00 atm, the temperature of the water is 10.0 °C, and the radius of the bubble is 5.00×10^{-3} m. At the surface, the pressure is 1.00 atm and the temperature of the water is 20.0 °C. (*a*) What is the final-to-initial volume ratio for the expanding bubble? (*b*) What is the change in thermal energy of the bubble? ••
51. A chamber fitted with a piston contains 1.00 mol of an ideal gas. (*a*) The piston is slowly moved to decrease the

chamber volume while the gas temperature is held constant at 296 K. How much work is done *by* the gas if the final chamber volume is one-third the initial volume? (*b*) If, instead of moving in the direction that decreases the chamber volume, the piston slowly moves in the direction that increases that volume, how much work is done *by* the gas if the final chamber volume is three times the initial volume? ●●

52. To change the volume of a gas isothermally by a factor f_1 of its initial value, it is necessary to do work W_1 on the gas. What work W_2 is needed to change the volume by a factor f_2 of the initial gas volume at the same temperature? Express your answer in terms of f_1, f_2, and W_1. ●●

53. An ideal gas initially at volume V_1, pressure P_1, and temperature T_1 undergoes an isobaric process that changes its temperature to T_2. The gas immediately undergoes an isothermal process that changes its volume to V_3. What is the final pressure in the gas in terms of these variables? ●●

54. At temperature T, a balloon is partly inflated with an ideal gas to a volume of 1.0 L. The balloon is then put in a low-pressure container at the same temperature T and its volume is increased to 2.0 L. If 26 J of work was done by the gas to accomplish this expansion, what is the final pressure inside the balloon in the low-pressure container? ●●

55. A 1.00-mol sample of an ideal gas expands such that its entropy doubles to 1.345×10^{24}. The temperature is held constant at 100 K during the expansion. (*a*) How much work is done by the gas on its surroundings? (*b*) What quantity Q of energy is transferred thermally to the gas? ●●

56. A 0.320-kg piston seals a chamber that contains 7.55×10^{24} helium atoms initially at 2.00 °C. The piston is free to move vertically, but it is originally held fixed in place with the chamber volume at 0.250 m^3. When the piston is released, the chamber volume immediately increases as the piston moves upward. At the instant when the piston has moved 0.320 m from its initial position, its velocity is 43.5 mm/s upward. The air pressure in the room is 1.00 atm, and the piston has a circular cross-section and a radius of 0.200 m. What is the final temperature of the gas if the expansion is (*a*) isentropic and (*b*) isothermal? ●●●

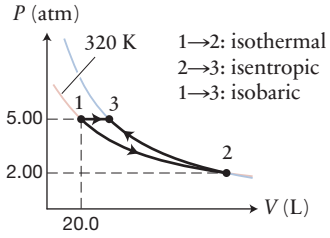
57. The ideal gas law is valid for a gas when, among other things, any interactions between particles are ignored. A more realistic model was proposed by Johannes Diderik van der Waals in 1873, for which he received a Nobel Prize in 1910. The van der Waals equation of state is

$$\left(P + \frac{an^2}{V^2}\right)(V - nb) = Nk_B T,$$

where a and b are constants, and n is the number of moles of the gas particles. Use this equation to determine the work done on a gas undergoing an isothermal process. Verify that your result reduces to the ideal gas case when $a = b = 0$. ●●●

58. A sample of diatomic ideal gas ($\gamma = 1.4$) initially has a volume of 20.0 L and is at 5.00 atm and 320 K. The gas undergoes the process shown in Figure P20.58. First it expands isothermally ($1 \rightarrow 2$) until the pressure is 2.00 atm, and then it is compressed isentropically ($2 \rightarrow 3$) until the pressure is 5.00 atm again. (*a*) Determine the volume of the gas after the isothermal expansion and the work done by the gas during the expansion. (*b*) Determine the volume and temperature of the gas after the isentropic compression, and the work done on the gas during the compression. (*c*) Determine the change in entropy of the gas by determining ΔS for $1 \rightarrow 2$ plus ΔS for $2 \rightarrow 3$. (*d*) Show that the entropy change is the same when the gas undergoes the isobaric process $1 \rightarrow 3$. ●●●

Figure P20.58



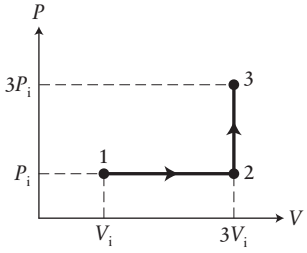
20.8 Entropy change in ideal gas processes

59. A sample of nitrogen gas ($\gamma = 1.4$) initially has a volume of 0.61 m^3 and a pressure of 1.00 atm. If the sample expands isentropically to a volume of 1.00 m^3, what is the final pressure? ●

60. A sample of an ideal gas for which $\gamma = \frac{7}{5}$ initially has a volume of 1.00 L and a pressure of 100 Pa. If the gas expands isentropically to 1.20 L, what is the final pressure? ●

61. Figure P20.61 shows the PV diagram for a monatomic ideal gas undergoing an isobaric expansion $1 \rightarrow 2$ followed by an isochoric pressure increase $2 \rightarrow 3$. If the sample contains N gas particles, what is the net change in entropy? ●●

Figure P20.61



62. An ideal gas for which $\gamma = 1.4$ initially has a volume of 1.5 m^3 and a pressure of 15 MPa. The gas is then slowly compressed isentropically to a volume of 0.50 m^3. (*a*) What is the final pressure in the gas? (*b*) By how much did the entropy of the gas change? ●●

63. A $\frac{1}{3}$-mol sample of an ideal gas is sealed in a container and heated from 273 K to 500 K. For this gas, what is the value of C_P in units of k_B if the temperature increase causes an entropy change of 1.8215×10^{23}? ●●

64. Two identical thermally insulated spherical tanks, A and B, are connected by a valve. Initially tank A contains 10 mol of an ideal diatomic gas, tank B is evacuated, and the valve is closed. If the valve is opened and the gas expands isothermally from A to B, what is the change in the entropy of the gas? ••

65. As part of a physics experiment, air under high pressure is allowed to expand isentropically out of a metal can and into an empty plastic bag. When the expansion is complete, the bag and contents are at atmospheric pressure. The air inside the can is initially at 5.0 atm and 25 °C. What is the air's temperature once it is in the plastic bag? ••

66. A gas is inside a cylinder fitted with a piston. The gas and its surroundings are both at 20 °C, and the gas is compressed as the piston moves and decreases the cylinder volume. The compression takes place slowly enough to allow the gas temperature to stay at 20 °C, and the work done on the gas by the compression is 2.0×10^3 J. What is the change in entropy of the gas? ••

67. An ideal gas that initially has a volume of 1.00 mL and is at 1.00 atm expands isentropically to 10.0 mL and 0.0215 atm. (*a*) How many degrees of freedom do the gas particles have? (*b*) What does this tell you about the structure of the particles? ••

68. A chamber fitted with a piston contains 1.00 mol of a monatomic ideal gas that is taken from an initial state i to a final state f via states 1 and 2. The chamber volume is initially $V_i = 0.100\ \text{m}^3$, and the temperature is $T_i = 273$ °C. To take the gas to state 1, it is heated isochorically to $T_1 = 373$ °C. To take it to state 2, it is heated isobarically to $T_2 = 473$ °C. Finally, the gas is allowed to expand isothermally until its volume is $V_f = 2V_2$. What is the change in entropy of the gas? ••

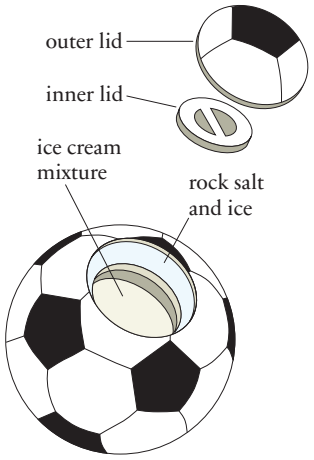
69. An ideal monatomic gas expands isentropically from an initial pressure P_i and initial volume V_i to a final pressure $P_f = P_i/20$ and final volume $V_f = 6V_i$. What is the work done by the gas, expressed in terms of P_i and V_i? [Hint: Use Eq. 20.8 (but remember that you want work done *by* the gas, so watch your signs) and the fact that $PV^\gamma =$ constant for an isentropic expansion (Eq. 20.46), with $\gamma = \frac{5}{3}$ for a monatomic ideal gas.] •••

20.9 Entropy change in nonideal gas processes

70. A thermodynamic process raises the temperature of 1.00 mol of water from 59.0 °C to 61.0 °C. What is the change in thermal energy of the water? •

71. An ice cream maker is in the form of a soccer ball (Figure P20.71). You fill the inner container with a mixture of cream and sugar at 32.0 °C and add rock salt and ice to the sleeve encircling the inner container. You then put both lids on and kick the ice cream maker around like a regular soccer ball until the internal temperature is low enough (−3.00 °C) to freeze the cream-sugar mixture into ice cream. (*a*) If the specific heat capacity of the mixture is $2900\,\text{J/K}\cdot\text{kg}$, what quantity of thermal energy must be removed from the inner container to bring 0.500 kg of the mixture to the temperature required for freezing? (*b*) If the mass of the maker plus its contents is 3.80 kg, how high above ground level would you need to lift the unit in order for the gravitational potential energy of the ice cream maker–Earth system to be equal to the quantity of energy you found in part *a*? ••

Figure P20.71



72. 2.00×10^4 J of energy is transferred thermally into a huge tank filled with liquid water. The water temperature remains constant at 10.0 °C during the process. By how much does the entropy of the water change? ••

73. An iceberg initially at 273 K is stopped in its motion by a small island. During the first few seconds of contact, 10 kJ of energy is transferred thermally from the island to the iceberg without the temperature of either object changing significantly. If the island temperature before the collision is 300 K, what is the change in entropy of the iceberg-island system as a result of the energy transfer? ••

Additional Problems

74. The average normal temperature of the human body is 98.6 °F. What is this temperature on the Celsius scale? •

75. Which are the most likely units for a measured temperature with a numerical value of (*a*) −322 and (*b*) 0? •

76. An ideal gas undergoes an isentropic process that changes its pressure from P_i to P_f and its volume from V_i to V_f. What is the work done on the gas? ••

77. A 22.7-g sample of diatomic oxygen gas molecules is in a sealed rigid container that has a volume of $0.0240\ \text{m}^3$. If the gas is initially at −1.50 °C, by how much does the pressure in the container change when the gas is cooled to −23.00 °C? ••

78. It is sometimes said, erroneously, that biological organisms violate the second law of thermodynamics because the biochemical reactions they carry out reduce their entropy by converting incoherent raw materials to coherent substances. To show that this is not so, consider your body at 37.0 °C transferring energy thermally into its immediate environment, where the temperature is 20 °C. For the system made up of your body and its environment, determine whether the entropy of the system increases or decreases over time, and at what rate. Assume your body transfers energy thermally at a rate of 100 W. ••

79. Two identical samples of a monatomic ideal gas start out at the triple point of water. Sample A is heated isochorically to five times its original temperature. Sample B is compressed isentropically and quasistatically to one-fifth of its original volume. In which process is the quantity of energy transferred thermally to the sample greater, and by how much? ●●●

80. You are working in a factory that uses water in its cooling systems and consequently has a lot of steam reservoirs. Due to carelessness, a 70.0-kg metal plate has been left on top of the only steam vent for a certain steam reservoir, and as a result the pressure in the reservoir builds up sufficiently to lift the plate a small distance. With the plate raised, the reservoir pressure drops slightly, and the plate lowers slowly back down and gently comes to rest on top of the vent. The vent is circular and has a diameter of 100 mm, and the plate is a square, 450 mm on a side. No steam is seen seeping out after this incident, and so you assume the plate has made a good seal such that only the circular section of plate directly above the vent is in contact with the steam. A flow meter attached to the reservoir tells you that 108 kg of liquid water was piped into that reservoir, and you know that this reservoir has a volume of 100 m^3. A coworker sees a few drops of water that have condensed on a glass portal located in the reservoir wall and thinks that the steam in the reservoir must have cooled enough to begin condensing to water (most of what's in the reservoir is still steam, though). Your coworker is about to grab the plate with his hands in order to uncover the vent and allow the steam to escape as it is designed to do when you suddenly realize that touching the plate might not be a good idea. ●●● CR

81. You are the founder of a start-up company that extracts electrical energy from ocean waves, a process called WaveGen. Your basic technology is a tube anchored vertically to the ocean floor. The unit is installed close to shore so that the bottom is anchored to the seabed and the top is above the water surface. At the top of the unit is a turbine generator for creating electricity. Waves rise and fall in the tube like a piston going up and down. As a wave rises in the tube, it pushes the air above it up into the generator. As the wave falls, air from the atmosphere flows into the tube through holes drilled in the part of the tube that is not under water. The average change in vertical wave height is 2.00 m, and the waves repeat on average every 5.00 s. A nearby seaside town requires 0.500 MW, and your goal is to supply all this electrical power. You know that the WaveGen tubes can be made economically up to a diameter of 2.00 m and that the efficiency of the generator in converting wave power to electricity is 68.0%, and you can produce a few tubes per day. ●●● CR



21

Degradation of Energy

CONCEPTS

21.1 Converting energy
21.2 Quality of energy
21.3 Heat engines and heat pumps
21.4 Thermodynamic cycles

QUANTITATIVE TOOLS

21.5 Entropy constraints on energy transfers
21.6 Heat engine performance
21.7 Carnot cycle
21.8 Brayton cycle

Some textbooks define energy as the *capacity to do work*. In the two preceding chapters, however, we saw that energy can do more than that. It can also be used to heat objects, and in this chapter we ask (and answer) two important questions about energy: To what extent can we use mechanical energy for heating and thermal energy for doing work? And, more specifically, to what extent can we convert energy from forms provided by nature to forms useful for technology and society?

The oceans of the world, for example, contain a very large amount of thermal energy. If we could convert just a small fraction of that energy to useful forms, we could satisfy the world's energy needs for thousands of years. Why can't a ship take in some of the thermal energy in the ocean and use it for propulsion? Such a scheme does not violate the energy law: Thermal energy would simply be converted to an equal amount of mechanical energy. As we see in this chapter, however, the entropy law imposes constraints on conversions of energy.

21.1 Converting energy

In Chapter 5 we saw that energy is a quantitative measure of the state of an object and that the law of conservation of energy implies that any change of state must be accompanied by a compensating change of state (or by a set of compensating changes of state). Whenever a pair of complementary changes of different types take place, energy is converted from one form to another. For example, burning propane gas involves two changes of state: a change in the molecular chemical bonds of the reagents (which we associate with chemical energy) and a change in the temperature of the gases (which we associate with thermal energy). The combustion is irreversible: It would be pretty startling if the hot gases would suddenly "unburn," cooling down and reconstituting the original propane and oxygen. Because the process is irreversible, we can conclude that the entropy of the gas increases in the process of burning (see Section 19.4).

A ball accelerating in free fall is another example of a complementary pair of changes of state: a change in the state of motion of the ball (which we associate with kinetic energy) and a change in the configuration of the Earth-ball system (which we associate with gravitational potential energy). This process is reversible: A movie clip of the falling ball played backward shows an upward moving ball slowing down as if it had been thrown upward (which *can* happen). We can thus conclude that the entropy of the Earth-ball system does not change.

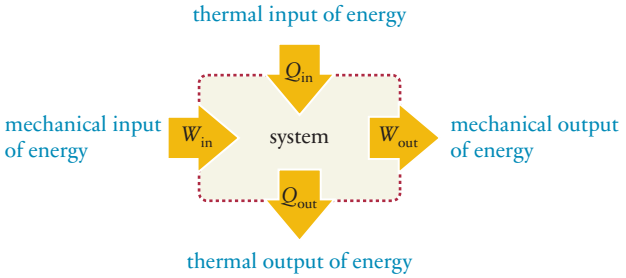
21.1 For each system, identify the complementary changes of state and associated energies, and determine whether or not the entropy of the system increases. (*a*) A puck slides to a stop on a surface. System: puck and surface. (*b*) A compressed spring is released and accelerates a cart on a low-friction track. System: spring and cart. (*c*) A projectile is fired from a cannon. System: projectile, cannon, and explosive chemicals. (*d*) A hot object and a cold object are placed in contact and reach thermal equilibrium. System: both objects.

To analyze the energy conversions that occur in a process, such as the ones in Checkpoint 21.1, we must choose a system and separately analyze all the energy transfers into and out of the system. As we saw in Chapter 20, energy can be transferred in two ways: mechanically and thermally. **Figure 21.1** shows for each of these two types of transfers an arrow representing the transfer of energy *into* the system (called *energy input*) and an arrow representing the transfer of energy *out of* the system (called *energy output*). Diagrams that show the energy inputs and energy outputs of a system are called **energy input-output diagrams.**

Energy inputs and energy outputs are always taken to be positive. The work done on a system is thus the difference between the **mechanical input of energy** and the **mechanical output of energy:** $W = W_{\text{in}} - W_{\text{out}}$. Likewise, the energy transferred thermally into a system is the difference between the **thermal input of energy** and the **thermal output of energy:** $Q = Q_{\text{in}} - Q_{\text{out}}$.

Figure 21.1 can represent the functioning of any device or industrial process. **Figure 21.2** on the next page shows the energy input-output diagrams for two devices: a car's internal combustion engine and a refrigerator. In a non-electric car, the combustion of gasoline provides a thermal input of energy to the engine (the system). The engine has two energy outputs: a mechanical output of energy to the wheels (and other parts) of the car and a thermal output of energy to the environment (Figure 21.2*a*). In a refrigerator there are two energy inputs (Figure 21.2*b*): a thermal input of energy from the refrigerator contents (which are part

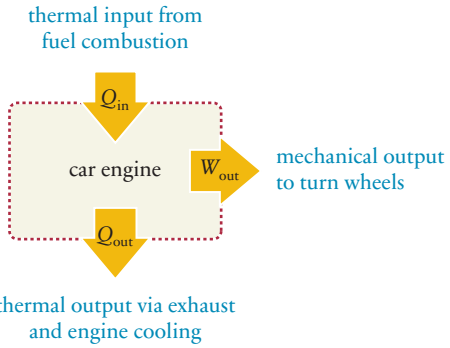
Figure 21.1 Energy input-output diagram.



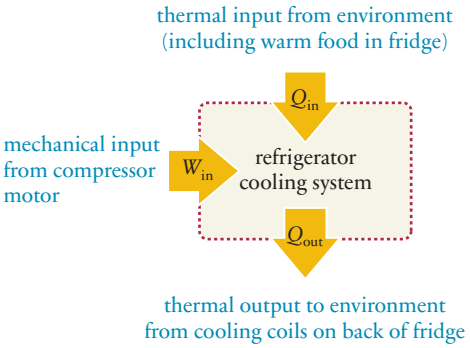
CONCEPTS

Figure 21.2 Examples of energy input-output diagrams.

(*a*) Energy input-output diagram for car engine



(*b*) Energy input-output diagram for refrigerator



of the environment) and a mechanical input of energy from the compressor motor to the rest of the refrigerator (the system). The refrigerator also has a thermal output of energy to the environment. (You may have noticed that the back of a refrigerator gets hot.)

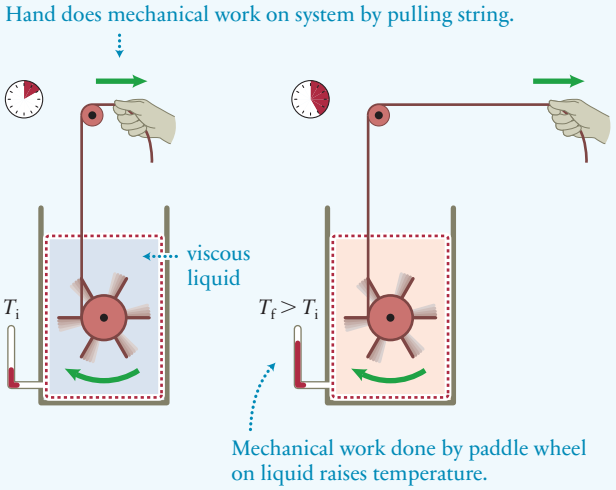
Each energy input or energy output in Figures 21.1 and 21.2 is accompanied by changes in entropy in the environment and in the system. To gain some practice determining energy and entropy changes, let us analyze a few specific situations.

CONCEPTS

EXERCISE 21.1 Converting mechanical energy to thermal energy

Consider the device illustrated in **Figure 21.3**. A string is attached to a paddle wheel that is immersed in a viscous liquid in a container. As the string is pulled, the paddle wheel rotates at constant speed, increasing the temperature of the liquid from T_i to T_f. (*a*) Draw an energy input-output diagram for the system comprising the paddle wheel and the liquid. Do the energy and entropy of the system comprising the paddle wheel and the liquid increase, decrease, or stay the same? (*b*) Assume the paddle wheel stops turning at the instant shown on the right in Figure 21.3, and we now wait until the liquid cools back

Figure 21.3

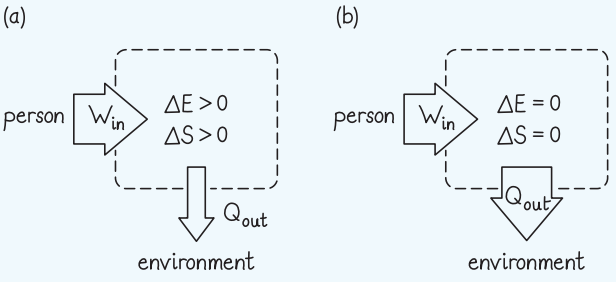


down to its initial temperature T_i. Draw an energy input-output diagram for the system, but this time over the time interval from the instant the wheel begins turning to the instant the liquid temperature is again T_i. Do the energy and entropy of the system increase, decrease, or stay the same over this new time interval?

SOLUTION (*a*) The work done on the system by pulling the string corresponds to an energy input. Because the system is not thermally insulated, energy is transferred thermally to the environment as the temperature of the system increases. My energy input-diagram therefore has one input and one output (**Figure 21.4*a***). Because the temperature of the system increases, I know that the energy of the system increases and so the energy input must be greater than the energy output. The increase in thermal energy increases the entropy of the system. ✔

(*b*) If the liquid cools back to its initial temperature, the initial and final states of the system are the same, so the energy change and the entropy change are both zero. The input of energy is the same as in part *a*, but because the energy of the system is not changing, the thermal output of energy must now be equal in magnitude to the mechanical input (Figure 21.4*b*). ✔

Figure 21.4



Note that the energy law imposes no restrictions on energy conversions: The energy law permits any form of energy to be converted entirely to any other form of energy. The device in Exercise 21.1, for example, converts mechanical energy entirely to thermal energy. Let us now look at the conversion of thermal energy to mechanical energy.

21.2 (*a*) For the situation illustrated in **Figure 21.5**, draw an energy input-output diagram for the gas. (*b*) Do the energy and entropy of the gas change? Ignore any friction in the piston and any energy transferred thermally out of the gas.

As you saw in Checkpoint 21.2, energy transferred thermally to the gas does work on the piston, and so the device converts thermal energy to mechanical energy. Because the temperature of the gas goes up, however, the thermal energy of the gas increases. Thus not all thermal energy is converted to mechanical energy in this device.

To prevent the thermal energy of the gas from increasing, we could immerse the setup in a thermal reservoir held at some fixed temperature T_R, as illustrated in **Figure 21.6**. If we then slowly remove grains of sand lying on top of the piston, the upward force exerted by the gas on the piston is a tiny bit greater than the downward force exerted by the remaining sand on the piston, and so the piston moves up. As the gas expands, it does work on the piston. Because the temperature of the gas remains constant, its thermal energy doesn't change, and so the energy required to raise the piston must be transferred thermally from the reservoir. Here is a situation where thermal energy is converted entirely to mechanical energy.

As a device, the setup of Figure 21.6 is not very practical because to maintain a constant gas temperature, the expansion would have to proceed *very* slowly. Moreover, in order for the gas to continue to do work on the piston, its volume would have to keep expanding.

Figure 21.5 When a volume of gas is heated at constant pressure in a cylinder fitted with a movable piston, both its temperature and its volume increase.

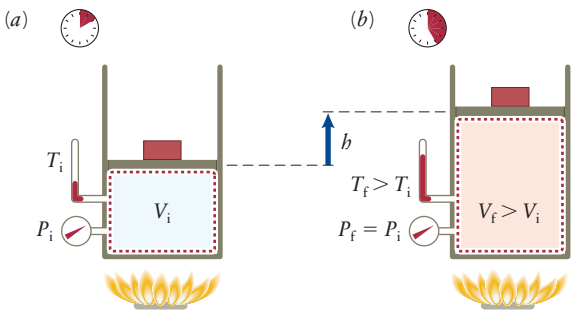
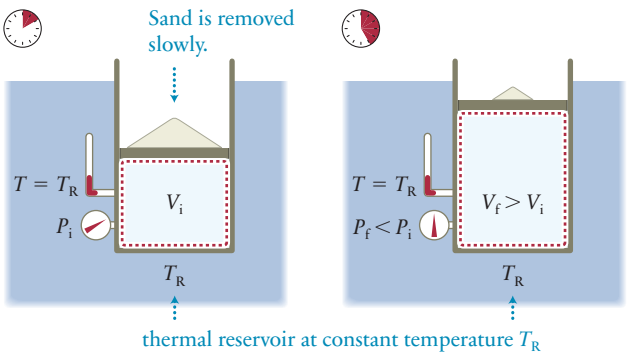


Figure 21.6 A gas held at constant temperature in a cylinder fitted with a movable piston expands slowly as the load on the piston is diminished.



When designing practical devices for converting energy, we usually want them to be able to operate indefinitely. If the device were to keep expanding in volume as in Figure 21.6 or keep getting hotter or building up pressure over time, it could not be operated for very long. For this reason we want the state of the device that converts energy to remain the same or to return to the same state over some repeating time interval. Such a device is called a **steady device.** When a steady device evolves from an initial state through a series of processes back to its initial state, the series of processes is called a *cycle*. During the operation of the device, this cycle is repeated continuously. For simplicity, we shall assume that all steady devices involve only quasistatic processes and that no energy is dissipated due to friction.

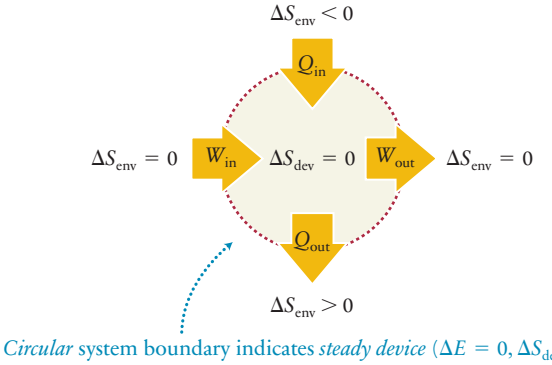
Because the initial and final states of a steady device are the same over one cycle, there is no change in the energy of the system: $\Delta E = 0$. Over one cycle, therefore, the energy input must be equal to the energy output: $W_{in} + Q_{in} = W_{out} + Q_{out}$.

To see what constraints are imposed by the entropy law (Eq. 19.5; $\Delta S > 0$ for a closed system moving toward equilibrium) on the energy inputs and outputs of a steady device, we examine the entropy changes associated with these energy inputs and outputs. The device is not a closed system and so we cannot apply the entropy law to it. The device and its environment constitute a closed system, however, so we can write $\Delta S = \Delta S_{dev} + \Delta S_{env} \geq 0$ (Eqs. 19.5, 19.6). Because the initial and final states of the cycle are the same, we know that there cannot be any change in the entropy of the device: $\Delta S_{dev} = 0$ for one cycle. Therefore, we need to consider only entropy changes in the environment when we deal with steady devices and make sure these changes satisfy $\Delta S_{env} \geq 0$.

As we saw in Section 20.1, no entropy change is associated with a quasistatic adiabatic process. Because $Q = 0$ for mechanical inputs and outputs of energy and because we assume that the processes in a steady device

CONCEPTS

Figure 21.7 Energy and entropy changes associated with energy inputs and energy outputs for a steady device.



are quasistatic, we have $\Delta S_{env} = 0$ for W_{in} and W_{out}. Therefore the only changes in the entropy of the environment we need to consider are those caused by thermal inputs and outputs of energy:

> **Steady devices operate under two constraints: (1) The energy input must be equal to the energy output. (2) The device's thermal inputs and outputs must increase the entropy of the environment (or leave it unaltered).**

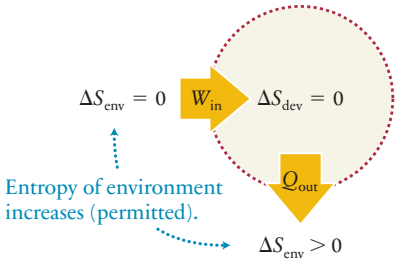
In Chapter 20, we saw that the entropy change associated with energy transferred thermally has the same sign as Q (Eq. 20.54). Therefore, a thermal output of energy increases the entropy of the environment. Conversely, a thermal input of energy decreases the entropy of the environment. **Figure 21.7** summarizes these entropy changes due to the various kinds of energy inputs and energy outputs for a steady device. We use a circular system boundary to indicate that the system is a steady device. The diagram represents the energy inputs and outputs during one complete cycle.

We are now in a position to analyze steady devices that convert energy. First consider a steady device that converts mechanical energy to thermal energy. Such a device has a mechanical input of energy and a thermal output of energy (**Figure 21.8a**). The mechanical input of energy involves no change in the entropy of the environment; the thermal output of energy causes an increase in the entropy of the environment. The entropy of the environment therefore increases, which is permitted by the entropy law. This means that a steady device can convert mechanical energy entirely to thermal energy.

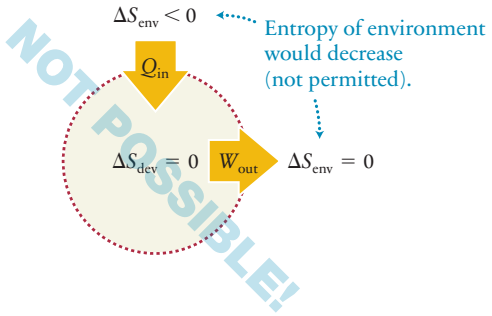
Next consider a steady device that converts thermal energy to mechanical energy (Figure 21.8*b*). The thermal input of energy causes a decrease in the entropy of the environment; the mechanical output of energy involves no change in the entropy of the environment. The entropy of the environment therefore decreases, which is not permitted by the entropy law for a steady device. Therefore no steady device can convert thermal energy entirely to mechanical energy.

Figure 21.8 Energy input-output diagrams for two types of steady device.

(*a*) Device that converts mechanical energy to thermal energy



(*b*) Attempt to convert thermal energy entirely to mechanical energy



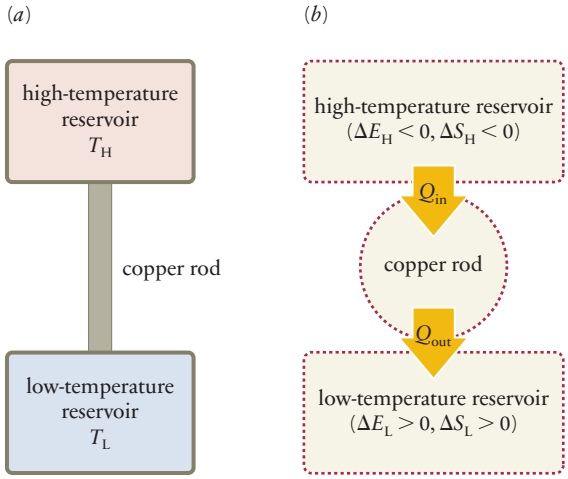
21.3 Suppose we slowly push the piston back to its original position in Figure 21.6, all the while maintaining a constant temperature in the gas. (*a*) Over the time interval that includes the expansion and the compression of the gas, has any energy been converted? (*b*) Does the entropy of the gas increase, stay the same, or decrease over that time interval? (*c*) If we keep repeating the expansion and compression, is this device a steady device?

21.2 Quality of energy

In this section we develop a graphical approach to keep track of entropy changes in the environment caused by steady devices that convert energy. Let us begin by considering the very simple situation illustrated in **Figure 21.9a**: A copper rod is in thermal contact with a high-temperature thermal reservoir on one end and a low-temperature thermal reservoir on the other end. Apart from its ends, the rod is thermally isolated from its environment. Because the temperatures of the thermal reservoirs are fixed, the rod remains in a steady state: One end remains at the temperature T_H of the high-temperature reservoir, and the other end remains at the temperature T_L of the low-temperature reservoir. (Remember that thermal reservoirs are large enough so that any energy transferred thermally to them does not change their temperature; see Section 20.7.) As a result of the temperature difference, energy is

CONCEPTS

Figure 21.9 (*a*) A copper rod in thermal contact with a high-temperature reservoir and a low-temperature reservoir. (*b*) Energy input-output diagrams for the rod and the reservoirs.



transferred thermally through the rod from the high-temperature reservoir to the low-temperature reservoir. Although the rod hardly qualifies as a device, its state doesn't change and so it satisfies the requirements that the entropy and energy of a steady device do not change over time. (The cycle for this device can be taken to be any arbitrary time interval.)

Figure 21.9*b* shows the energy input-output diagrams for the rod and for the two thermal reservoirs. The change in the energy of the high-temperature reservoir is equal to the thermal input of energy of the rod, $\Delta E_H = -Q_{in}$, while the change in the energy of the low-temperature reservoir is equal to the thermal output of energy of the rod, $\Delta E_L = Q_{out}$. Because the state of the rod is not changing, we know that $\Delta E_{rod} = Q_{in} - Q_{out} = 0$, so $Q_{in} = Q_{out}$ and therefore $\Delta E_H = -\Delta E_L$. The amount of energy transferred out of the high-temperature reservoir is equal to the amount of energy transferred into the low-temperature reservoir.

The thermal transfer of energy through the rod causes entropy changes in the two reservoirs (both part of the rod's environment) even though their temperatures remain unchanged. Because energy is transferred out of the high-temperature reservoir, its energy decreases, and so its entropy decreases too: $\Delta S_H < 0$. The increase of energy in the low-temperature reservoir causes its entropy to increase: $\Delta S_L > 0$. Because the transfer of energy is irreversible, we know that the entropy of the environment must increase, so that $\Delta S_H + \Delta S_L > 0$.

The reason entropy increases can be understood by looking at the way we defined temperature in Chapter 19: $dS/dE \equiv 1/(k_B T)$ (Eq. 19.38).* The quantity dS/dE, called the **entropy gradient**, is a measure of the change in entropy caused by a given change in energy. Equation 19.38 tells us that the magnitude of the change in entropy caused by transferring a given amount of energy to or from a reservoir is inversely proportional to the temperature of the reservoir. The entropy change ΔS caused by a given change in energy ΔE is therefore greater in magnitude for the low-temperature reservoir than for the high-temperature reservoir. Because ΔS_L is positive and ΔS_H is negative, it follows that $\Delta S_H + \Delta S_L > 0$.

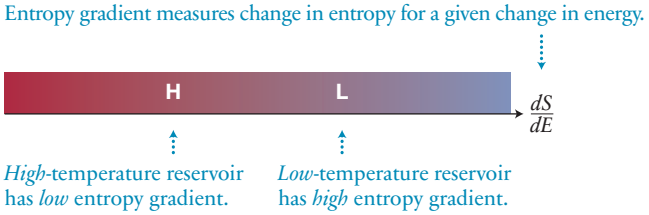
21.4 What are the signs of ΔE_R, ΔS_R, and $\Delta S_R/\Delta E_R$ for a thermal transfer of energy (*a*) to and (*b*) from a thermal reservoir?

You can view the entropy gradient as the "entropy cost of an energy change." The greater the entropy gradient (or the lower the temperature), the greater the change in entropy caused by a given energy change. As you saw in Checkpoint 21.4, the entropy gradient is always positive, regardless of the direction of the energy transfer, because ΔS and ΔE always have the same algebraic sign.

We can arrange the rod's environment by its entropy gradient. Because the high-temperature reservoir has the smaller entropy gradient, we place it to the left of the low-temperature reservoir on the entropy-gradient scale in **Figure 21.10**. If we represent the thermal transfer of energy through the rod in **Figure 21.11a** on the next page, by a curved arrow, we obtain the diagram shown in Figure 21.11*b*. The tail of the arrow represents the thermal input of energy from the high-temperature reservoir, and the head represents the thermal output of energy to the low-temperature reservoir.

Note that we can associate the sign of the entropy changes in the reservoirs with the direction of the arrow. The tail of the arrow (the energy input) points down and $\Delta S_H < 0$; the head (the energy output) points up and $\Delta S_L > 0$. The diagram in Figure 21.11*b* therefore allows us to see that the transfer of energy through the rod increases the entropy of the environment: The rod takes a certain amount of energy out of the environment at low entropy cost (small dS/dE) and returns the same

Figure 21.10 Entropy gradient. The shaded bar represents the rod's environment, which consists of the high-temperature reservoir (H) and the low-temperature reservoir (L).

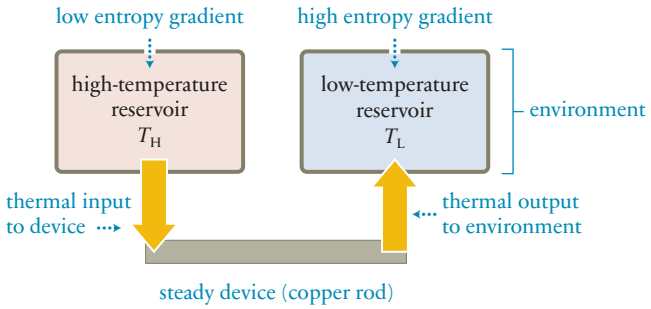


*As we did in Chapter 20, we assume that the kinetic, potential, and source energies of the reservoirs remain constant, so that $\Delta E = \Delta E_{th}$.

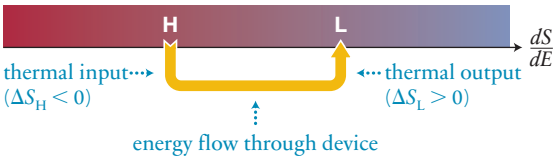
CONCEPTS

Figure 21.11 (*a*) We rearrange the device of Figure 21.9 to separate the system (the steady device) from its environment (the thermal reservoirs). (*b*) Using this geometry, we construct an entropy diagram showing the energy conversions and transfers in the steady device. The thermal input at low dS/dE causes entropy change in the environment of smaller magnitude than thermal output at high dS/dE, so the entropy of the environment increases: $\Delta S_H + \Delta S_L > 0$.

(*a*) We separate the device (the system) from its environment



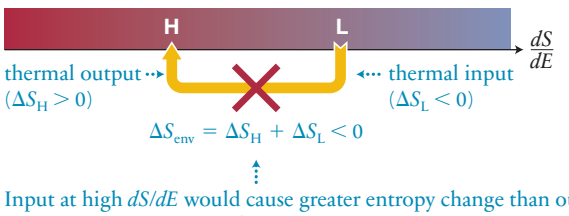
(*b*) Entropy diagram



amount of energy to the environment at a greater entropy cost (large dS/dE). The magnitude of the entropy change caused by the output is therefore greater than the magnitude of the entropy change caused by the input. Because the input causes a decrease in entropy and the output causes an increase in entropy, the combined effect is an increase in the entropy of the environment.

Now suppose the copper rod somehow transfers energy from the low-temperature reservoir to the high-temperature reservoir, as illustrated in **Figure 21.12**. The curved arrow representing the energy transfer now points from the low-temperature reservoir to the high-temperature reservoir and, as you can readily verify, the signs of the entropy changes are the opposite of the signs in Figure 21.11. The rod takes a certain amount of energy out of the environment at large dS/dE and returns the same amount of energy to the environment at a smaller dS/dE. The entropy change of the environment in this process is thus negative, which is not allowed by the entropy law. Indeed, no one has ever observed energy being transferred thermally from a cold object to a hot object with no other changes in the system or the environment (that is, the cold object spontaneously getting colder and the hot object getting hotter).

Figure 21.12 Entropy diagram for a thermal transfer of energy from a low-temperature reservoir (L) to a high-temperature reservoir (H), a process forbidden by the entropy law.



The diagrams in Figures 21.11 and 21.12 are called **entropy diagrams.** An entropy diagram shows the energy conversions and transfers that take place in a steady device with the environment arranged according to an entropy gradient, as shown in Figure 21.10. Arrows in the diagram represent the energy transfers through the device during one complete cycle. The tails of the arrows represent energy inputs to the device, and the heads represent energy outputs. Thermal inputs of energy (tails) cause the entropy of the environment to decrease, while thermal outputs of energy (arrowheads) cause the entropy of the environment to increase.

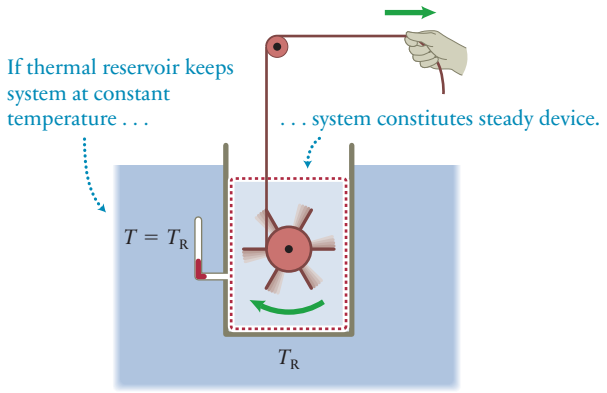
Entropy diagrams permit us to keep track of entropy changes that result from the energy conversions by a steady device. In particular:

> **A transfer of energy that points to the right in an entropy diagram causes an increase in the entropy of the environment; a transfer that points to the left causes a decrease in the entropy of the environment.**

As we shall see in Section 21.5, entropy diagrams can be used to derive quantitative information about entropy changes in the environment.

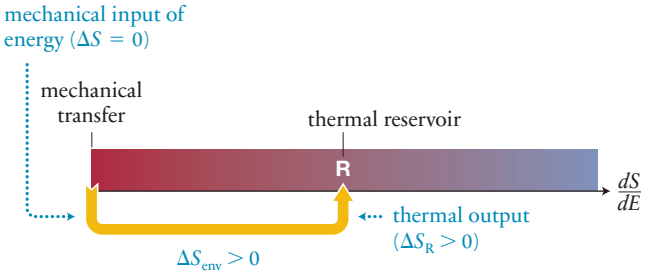
Let's now return to the example of a paddle wheel in a viscous liquid. This time, however, we immerse the container of viscous liquid in a large thermal reservoir that keeps the viscous liquid at a constant temperature T_R (**Figure 21.13**). The moving parts and the viscous liquid now constitute a steady device. A person does work on the device by pulling on the string and then the device transfers energy to the reservoir. The device thus converts mechanical energy to thermal energy.

Figure 21.13 Paddle-wheel device from Exercise 21.1 is held at a constant temperature T_R by a large thermal reservoir of water.



CONCEPTS

Figure 21.14 Entropy diagram for the process depicted in Figure 21.13.



To draw an entropy diagram for this energy conversion process, we must first determine where to put mechanical energy on the entropy-gradient scale. Remember that doing work—in other words, transferring energy mechanically—does not cause any change in entropy. Therefore $\Delta S = 0$ and so $\Delta S/\Delta E = 0$. The entropy gradient associated with mechanical energy is thus zero. Mechanical energy transfers therefore always appear on the far left at the origin of the dS/dE scale (**Figure 21.14**).

The only entropy change in Figure 21.14 is thus caused by the thermal output of energy. This entropy change is positive (arrow points up), and so the entropy of the environment increases.

The inverse of the process in Figure 21.14, illustrated in **Figure 21.15**—conversion of thermal energy entirely to mechanical energy by a steady device—cannot occur. In this process, the device would transfer energy thermally from the environment and mechanically to the environment. The curved arrow now points to the left, causing a decrease in the entropy of the environment, which is impossible. This process would mean that the paddle wheel in Figure 21.13 could rewind the string and pull the hand to the left by transferring energy thermally from the reservoir. The entropy law prevents this from happening.

Summarizing, we see that converting energy with a low entropy gradient to energy with a high entropy gradient with a steady device (arrow points right in entropy diagram) increases the entropy of the environment and is permitted by the entropy law. Conversely, converting

Figure 21.15 Entropy diagram for the conversion of thermal energy entirely to mechanical energy by a steady device.

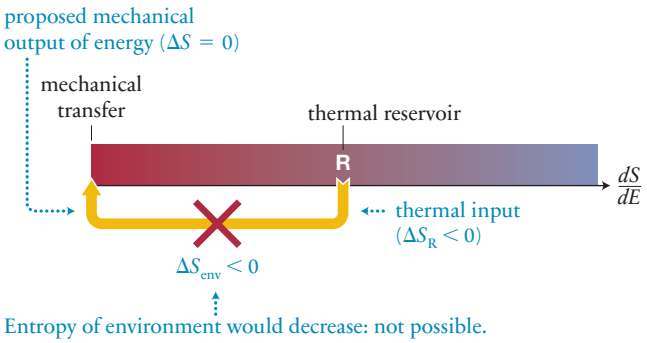
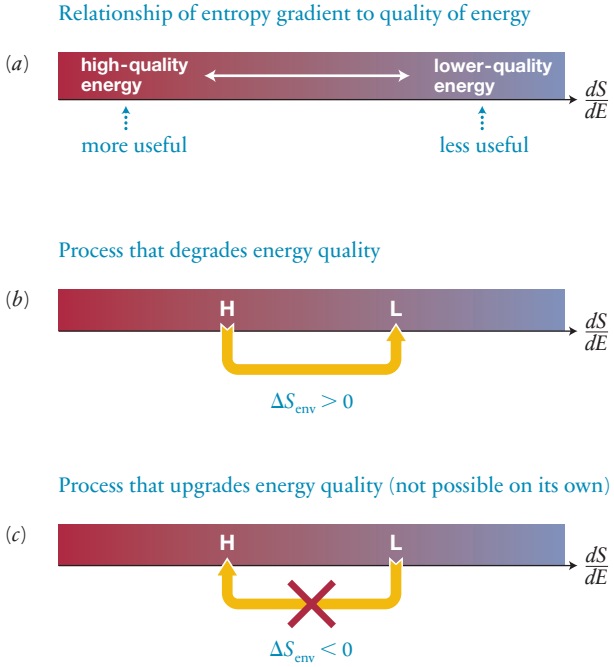


Figure 21.16 Energy with a large entropy gradient is less useful and therefore of lower quality than energy with a small entropy gradient.



energy with a high entropy gradient to energy with a low entropy gradient with a steady device (arrow points left in entropy diagram) decreases the entropy of the environment and is not permitted by the entropy law.

As energy shifts toward the right in an entropy diagram, the energy becomes less useful for, say, heating things or powering machines. You can do less with a tub of lukewarm water than with a bucket of boiling hot water that contains the same amount of thermal energy as the tub. We can therefore associate the entropy gradient with the **quality of the energy:** The smaller dS/dE, the lower the entropy cost and the higher the quality of energy (**Figure 21.16a**). The tendency of the entropy of the environment to increase and energy to shift from higher quality to lower quality (that is, to the right in the entropy diagram) corresponds to a *degradation* of energy quality (Figure 21.16*b*). In the device of Figure 21.13, for instance, high-quality mechanical energy (which allows us to do useful work, such as lifting or accelerating things) is irreversibly turned into less useful, lower-quality thermal energy.

A conversion of lower-quality energy to higher-quality energy (an *upgrade* in energy quality) corresponds to a left-pointing curved arrow in entropy diagrams and produces a decrease in the entropy of the environment (Figure 21.16*c*). Such upgrades cannot occur in steady devices that have a single energy input and a single energy output. For example, it's impossible to convert low-quality thermal energy entirely to high-quality mechanical energy. This is why ships can't be powered by converting thermal energy from the ocean entirely to mechanical energy.

21.5 For a steady device that has a single input of energy and a single output of energy, does the quality of the energy increase, decrease, or stay the same when the entropy change in the environment is (*a*) positive and (*b*) zero?

EXAMPLE 21.2 Energy and entropy changes in a thermal reservoir

Suppose the person pulling the string in Figure 21.13 does work W on the device immersed in a thermal reservoir R at temperature T_R. (*a*) Do the energy and entropy of the reservoir increase, decrease, or stay the same? (*b*) If the same experiment is repeated with the reservoir at a lower temperature, are the changes in the energy and entropy of the reservoir greater than, the same as, or smaller than in part *a*?

1 GETTING STARTED Because the device is immersed in a thermal reservoir, the state of the device doesn't change. The device is therefore a steady device, and its energy and entropy don't change. Because the reservoir temperature is lower in part *b*, the states of the device and reservoir are different from those in part *a*. In particular, a lower reservoir temperature means that the reservoir's thermal energy is smaller.

2 DEVISE PLAN I can draw an energy input-output diagram for the device and determine its energy inputs and outputs. From the diagram, I can deduce whether there is any thermal transfer of energy to (or from) the reservoir. To evaluate any entropy changes in the reservoir, I draw an entropy diagram for each part of the problem.

3 EXECUTE PLAN (*a*) I know that the person does work on the device and so $W_{in} = W > 0$. Because the energy of the device doesn't change, the energy law tells me that $\Delta E = W + Q = 0$, so $Q = -W < 0$, which means that energy must be transferred thermally out of the device: $Q_{out} = -Q = W$. My energy input-output diagram is shown in **Figure 21.17*a***.

The energy transferred thermally out of the device goes into the reservoir, which tells me that the energy of the reservoir increases by an amount W. ✔

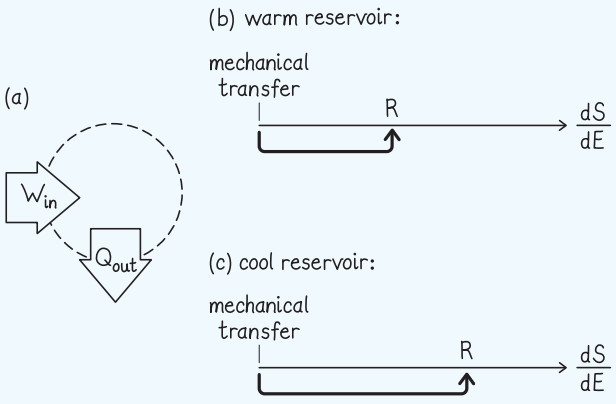
To see what happens to the reservoir entropy, I need an entropy diagram. Mechanical energy initially in the device is converted to thermal energy. I represent this conversion by drawing a right-pointing arrow in my entropy diagram, running from the origin ($dS/dE = 0$ for mechanical energy) to the R representing the reservoir (Figure 21.17*b*). That the arrow points to the right tells me that the entropy of the reservoir increases. ✔

(*b*) Even though the temperature of the reservoir is lowered, the amount of work done on the device by the person is still W, so the temperature change does not affect my energy input-output diagram. This tells me that the energy of the reservoir increases by the same amount W as in part *a*. ✔

Because the reservoir temperature is lower than in part *a*, however, on my entropy diagram for this situation (Figure 21.17*c*), I show R farther to the right than in Figure 21.17*b*. The right-pointing arrow representing the conversion is longer, meaning a greater dS/dE value at this temperature. A greater entropy gradient means a higher entropy cost of energy change, and so the entropy change associated with the energy change W must be greater than the entropy change of part *a*. ✔

4 EVALUATE RESULT I can also evaluate the entropy changes using Eq. 20.54, $\Delta S_R = Q_R/k_B T_R$. The quantity of energy transferred thermally to the reservoir is the same for both parts of the problem: $Q_R = W$. Because the reservoir temperature T_R is lower in part *b*, the entropy change must be greater here than in part *a*.

Figure 21.17



21.6 In an entropy diagram, which of these statements are true? (*a*) The quality of the energy is proportional to the entropy gradient. (*b*) The quality of the energy is inversely proportional to the entropy gradient. (*c*) High-temperature energy is of high quality. (*d*) High-temperature energy is of low quality. (*e*) Low-temperature energy has a small entropy gradient. (*f*) Low-temperature energy has a large entropy gradient.

21.3 Heat engines and heat pumps

As we just saw, it is not possible for a steady device that has a single energy input and a single energy output to upgrade energy quality. The only way for a steady device to upgrade energy quality is to also degrade energy quality in such a way that the change in entropy of the environment that results from both processes is either positive or zero: $\Delta S_{env} = \Delta S_{upgrade} + \Delta S_{degrade} \geq 0$.

The process illustrated in **Figure 21.18** is one way to combine two compensating processes. A steady device takes in energy from a high-temperature reservoir, converts part of this energy to mechanical energy, and thermally transfers the rest to a low-temperature reservoir. This process is allowed as long as the increase in entropy associated with the degradation of the energy (the thermal output) equals or exceeds the decrease in entropy associated with the upgrade (the mechanical output). A steady device of this type is called a **heat engine.**

CONCEPTS

Figure 21.18 Entropy diagram for a process that upgrades and degrades energy.

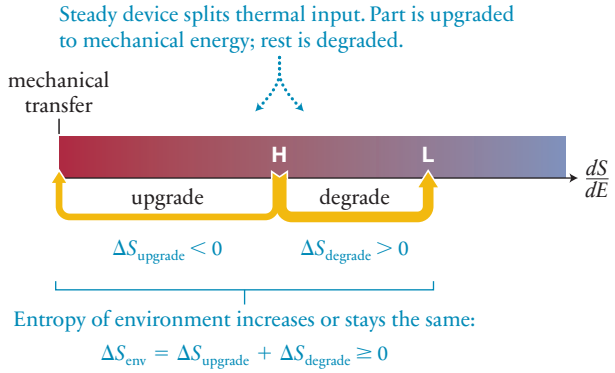
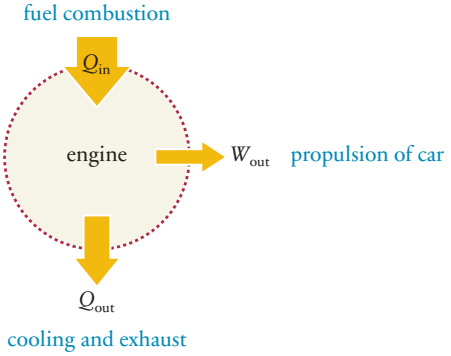


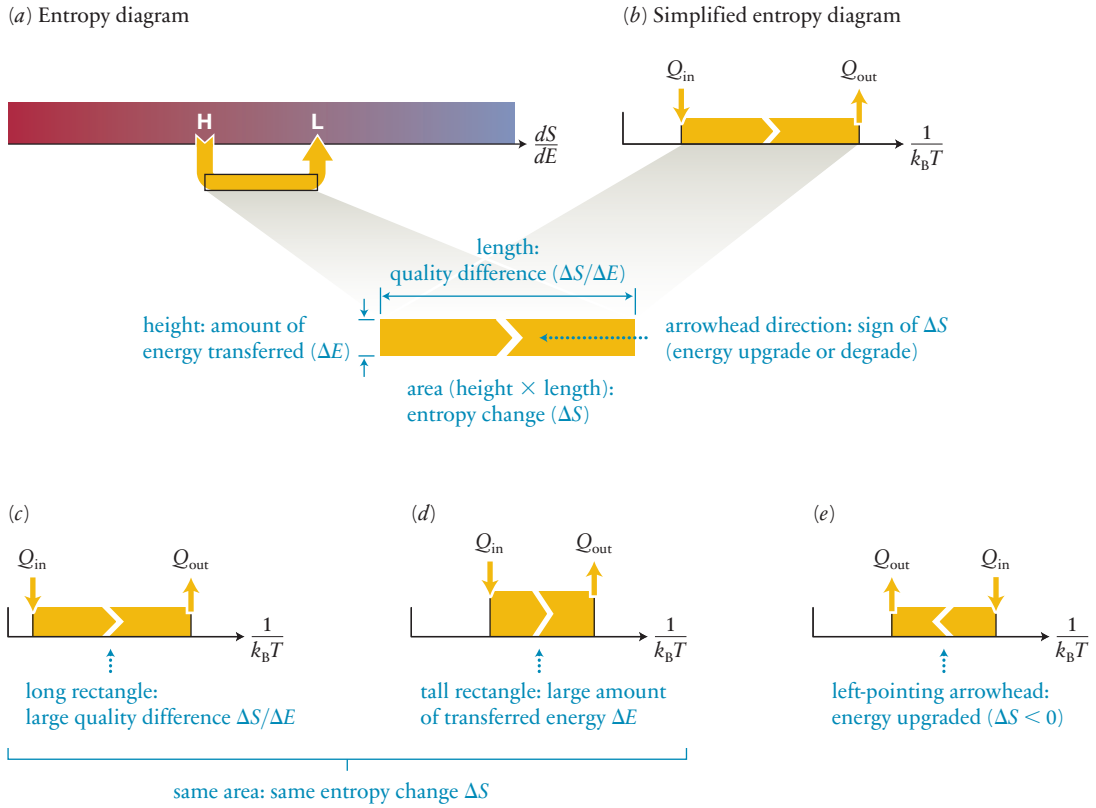
Figure 21.19 Energy input-output diagram for an automobile internal combustion engine, which is one type of heat engine.



One familiar heat engine that carries out the process illustrated in Figure 21.18 is the internal combustion engine we discussed briefly in Section 21.1. As shown in **Figure 21.19**, the thermal input of energy is provided by the combustion of gasoline at high temperature (about 2700 K). Roughly two-thirds of this input energy is discarded to the environment by the car's exhaust and cooling system, both of which are roughly at 400 K. The remaining one-third of the input energy is turned into mechanical output of energy (part of which is lost due to friction in the moving parts of the car). As you can see by the relative sizes of the Q_{out} and W_{out} arrows in Figure 21.19, in a typical internal combustion engine more energy is degraded than upgraded in order to satisfy the entropy law and keep the engine's final state the same as its initial state.

Because we arrange energy inputs and outputs on the horizontal scale of an entropy diagram according to their entropy gradient, the length of an arrow in the diagram is a measure of the quality difference $\Delta S/\Delta E$ between the input and output. If we take the width of each arrow to represent the amount of energy ΔE transferred, then the area of the horizontal part of the arrow (the product of the arrow's width and length) represents the magnitude of the change in entropy (**Figure 21.20*a***).

Figure 21.20 Obtaining entropy changes from entropy diagrams.



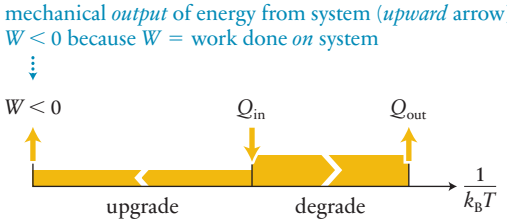
CONCEPTS

Because it is difficult to draw curved arrows and determine their area, we simplify our procedure for drawing entropy diagrams, replacing the arrows with rectangles (Figure 21.20*b*). The length of the rectangle corresponds to the quality difference between the energy output and input, and the height of the rectangle corresponds to the amount of energy transferred. The area of the rectangle thus represents the entropy change in the environment caused by the energy conversion or transfer. The direction of the energy transfer is indicated by an arrow inside the rectangle. In addition we mark the energy inputs and outputs by vertical arrows: down for input and up for output. Finally, because the entropy gradient dS/dE is equal to $1/k_BT$ (Eq. 19.38), we label the axis with $1/k_BT$.

There are two ways to increase the magnitude of the entropy change. The greater the quality difference between the input and output, the greater the entropy change. For example, in Figure 21.20*c* the energy input is of higher quality than in Figure 21.20*b*. Consequently, the entropy change produced by the conversion of this higher-quality energy is greater. Shifting the input of the arrow to the left (higher quality) increases its length and therefore its area (which represents the entropy change). As you can see from Eq. 20.54, $\Delta S = Q/k_BT$, the entropy change is also proportional to the quantity of energy Q being transferred. So, a second way to induce a greater entropy change between two fixed-temperature reservoirs is to increase the amount of energy being transferred (represented by the greater height of the rectangle in Figure 21.20*d*). Finally, if the direction of an energy transfer is reversed, the sign of the entropy change is reversed but its magnitude remains the same (compare Figure 21.20*e* with 21.20*b*). **Figure 21.21** shows the entropy diagram for a process that both upgrades and degrades energy, with rectangles used to represent the transfers.

Because thermal inputs and outputs of energy occur at different temperatures, they are separated on the dS/dE scale. Mechanical inputs and outputs of energy, however, are always located at $dS/dE = 0$, and so they overlap in an entropy diagram. The end of an arrow (or rectangle) at $dS/dE = 0$ therefore does not represent

Figure 21.21 Entropy diagram for a process that upgrades and degrades energy (compare with Figure 21.18).

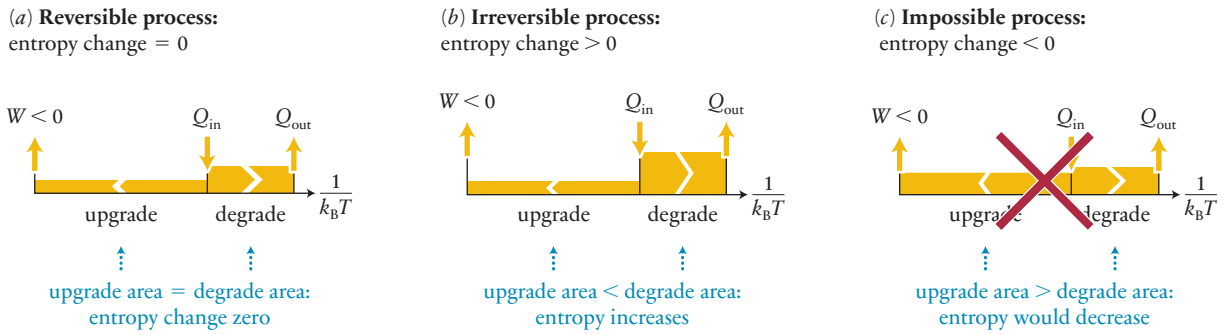


W_{in} or W_{out} but rather the difference between these two, $W = W_{in} - W_{out}$, which is the work done on the system. When the arrow points up, energy is transferred mechanically from the system to the environment and hence $W < 0$ (Figure 21.21). When the arrow points down, work is done on the system, indicating a mechanical transfer of energy into the system, $W > 0$.

For a steady device, the entropy law requires the entropy change in the environment to be greater than or equal to zero, and so the area of the degrade rectangle must be larger than or equal to the area of the upgrade rectangle. **Figure 21.22*a*** shows a process for which the energy upgrade and energy degradation cause equal-magnitude changes in entropy. For such a process, the entropy change is zero, so the process is reversible and allowed. Figure 21.22*b* shows a process for which the magnitude of the entropy change caused by the energy upgrade is smaller than the magnitude of the entropy change caused by the energy degradation. For such a process, the entropy change is positive, so the process is irreversible and allowed. Figure 21.22*c* shows a process for which the magnitude of the entropy change caused by the energy upgrade is greater than the magnitude of the entropy change caused by the energy degradation. For such a process, the entropy change is negative, so the process is forbidden by the entropy law. The procedure for drawing entropy diagrams is shown in the Procedure box on page 754.

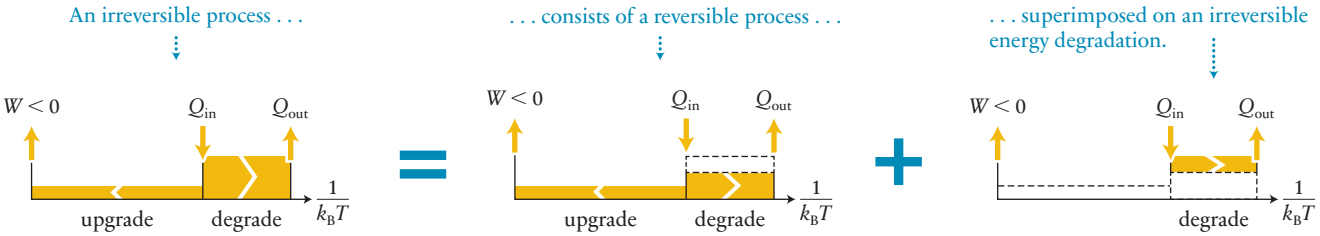
21.7 Draw an entropy diagram for the refrigerator in Figure 21.2*b*.

Figure 21.22 For a steady device, the entropy change in the environment must be greater than or equal to zero.



CONCEPTS

Figure 21.23 Entropy diagram for an irreversible process.



The entropy diagram for an irreversible process such as the one in Figure 21.22*b* can be treated as the entropy diagram for an irreversible degradation of energy (that is, hot to cold) superimposed on the entropy diagram for a reversible process (**Figure 21.23**). The upgrade rectangles in Figure 21.23*a* and *b* are the same, but in Figure 21.23*b* the height of the degrade rectangle has been adjusted so that the areas of the upgrade and degrade rectangles are the same. The entropy change in Figure 21.23*b* is therefore zero, and the process represented by this entropy diagram is reversible. Figure 21.23*c* shows the remainder of the energy transferred to the low-temperature reservoir. This process increases the entropy of the environment and is irreversible. The mechanical energy generated in Figure 21.23*b* is retrievable. Because the process is reversible, the mechanical energy can be converted back to thermal energy. The energy in Figure 21.23*c*, on the other hand, has been degraded and is irretrievable. It cannot be transferred back to the high-temperature reservoir without some other associated energy degradation. For now, the only processes we consider are reversible ones for which the entropy change is zero.

21.8 Why must the degrade rectangle in Figure 21.22*a* have a greater height than the upgrade rectangle?

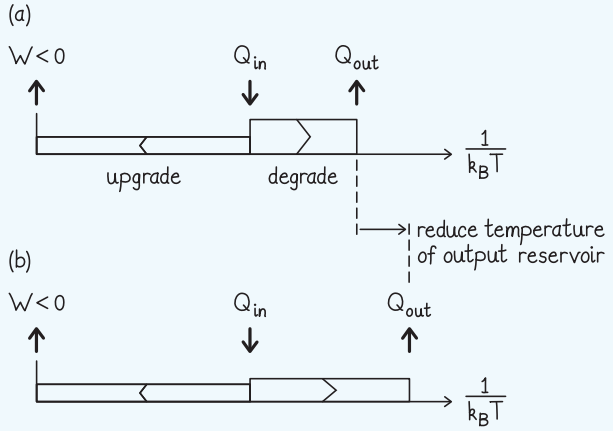
EXAMPLE 21.3 Hot or cold?

A reversible heat engine converts energy taken from a thermal reservoir at temperature T_{in} by doing work on the environment and thermally transferring energy to a thermal reservoir at temperature $T_{out} < T_{in}$. How do the thermal input and output of energy change if the temperature of the output reservoir is lowered without changing the amount of work done by the engine on the environment?

1 GETTING STARTED I begin by drawing an entropy diagram for the heat engine (**Figure 21.24*a***). Because the engine is reversible, the area of the upgrade rectangle must equal the area of the degrade rectangle. The height of the upgrade rectangle is equal to the magnitude of W and is fixed. The height of the degrade rectangle is equal to Q_{out}.

2 DEVISE PLAN To determine how the thermal input and output of energy change as the temperature of the output reservoir is lowered, I need to draw another entropy diagram (Figure 21.24*b*), keeping the upgrade rectangle the same as in Figure 21.24*a* and shifting the right side of the degrade rectangle to the right to represent the lower temperature of the output reservoir. When I make this change, I must remember that the upgrade and degrade rectangles must still have the same area, and so now my degrade rectangle is not as tall as in Figure 21.24*a*. Also, I know that the sum of the outputs must equal the sum of the inputs, and so $W_{out} + Q_{out} = Q_{in}$.

Figure 21.24 Example 21.3.



3 EXECUTE PLAN In any entropy diagram, the height of the degrade rectangle represents Q_{out}. Because the height of the degrade rectangle is less in Figure 21.24*b* than in Figure 21.24*a*, I conclude that Q_{out} decreases when the temperature of the output reservoir is lowered. Because $W_{out} + Q_{out} = Q_{in}$ and W_{out} is not changed, Q_{in} must also decrease. ✔

4 EVALUATE RESULT My result means that if the temperature difference between the reservoirs increases, I need less energy Q_{in} to do the same amount of work on the environment. The reason is that I can compensate for the energy upgrade from thermal to mechanical with less energy: At low temperature a given amount of energy causes a greater entropy change and so my answer makes sense.

21.9 Suppose you keep W and T_{out} in the reversible process of Figure 21.24*a* fixed but increase T_{in}. Do the magnitudes of these quantities increase, decrease, or stay the same: (*a*) ΔS associated with the upgrade, (*b*) ΔS associated with the degradation, (*c*) Q_{in}?

CONCEPTS

PROCEDURE: Drawing entropy diagrams for a steady device

1. Draw an entropy gradient axis pointing right. Add a short vertical axis to indicate the origin of the axis. Label the axis $1/k_BT$ to remind yourself of the relationship between entropy gradient and temperature.
2. Mark the positions along the axis of the thermal inputs and outputs of energy of the device. Transfers at higher temperature are on the left; transfers at lower temperature are on the right. Use downward-pointing arrows for inputs, use upward-pointing arrows for outputs, and label these Q_{in} and Q_{out}.
3. Mechanical transfers of energy are placed at the origin of the axis. Use a downward-pointing arrow when the work done on the system is positive and an upward-pointing arrow when the work done on the system is negative. Label this arrow by W and indicate whether $W > 0$ or $W < 0$.
4. Draw rectangles to represent each energy transfer. Make the height of each rectangle proportional to the amount of energy transferred (but see point 5 below). Use an arrowhead to indicate the direction of each transfer.
5. For the operation of the device to be permitted by the entropy law, the combined area of the rectangles in which the arrows point right must be at least as large as the combined area of the rectangles in which the arrows point left.

Example 21.3 and Checkpoint 21.9 show that it is beneficial to take energy in at the highest possible temperature and discard energy at the lowest possible temperature. Doing so minimizes the thermal input of energy for a fixed amount of work done on the environment.

We can define the **efficiency** of a heat engine as follows:

> **The efficiency of a heat engine is the ratio of the work done by the engine on the environment to the thermal input of energy into the engine. For a heat engine operating between two fixed-temperature reservoirs, the efficiency is greatest when the thermal input of energy takes place at the highest possible temperature and the thermal output of energy takes place at the lowest possible temperature.**

We see now that no heat engine—even a reversible one—can have an efficiency of 1.00 because some fraction of the thermal input of energy has to be degraded to make up for the entropy decrease that occurs in the process of upgrading the energy. The imperfect efficiency of heat engines has nothing to do with friction in the moving parts or other irreversible phenomena; rather, it is a direct consequence of the entropy law.

CONCEPTS

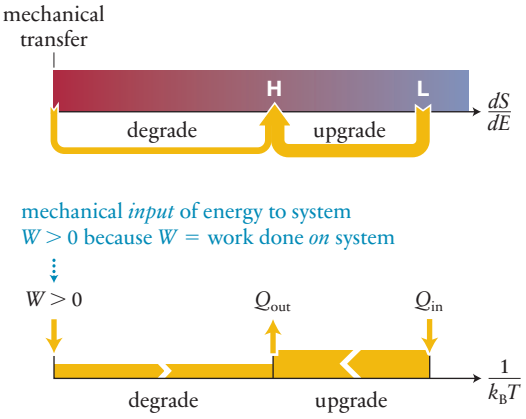
21.10 When the temperature of the output reservoir in Example 21.3 is lowered, does the efficiency increase, remain the same, or decrease?

As you saw in Checkpoint 21.7, there are more ways to cause a degradation of energy that compensates for an upgrade than the one shown in Figure 21.21. For example, you can swap the roles of the upgrade and degradation shown in Figure 21.21. The resulting entropy diagram, shown in **Figure 21.25a**, represents the operation of a **heat pump.** In a heat pump, mechanical energy is degraded in order to thermally transfer energy from a lower-temperature region of the environment to a higher-temperature region of the environment—a direction opposite the natural direction in which energy is transferred thermally. Note that work is done on the device, so $W > 0$.

The refrigerator is one familiar example of a heat pump (Figure 21.25*b*). The interior of the refrigerator constitutes the reservoir at T_{in}; the air surrounding the

Figure 21.25 Energy conversions in a heat pump.

(*a*) Entropy diagrams for energy conversions in heat pump



(*b*) Energy input-output diagram for refrigerator (a type of heat pump)

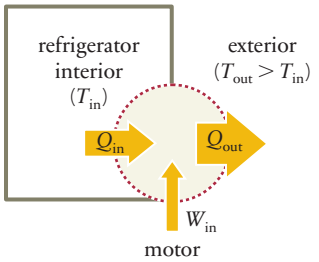
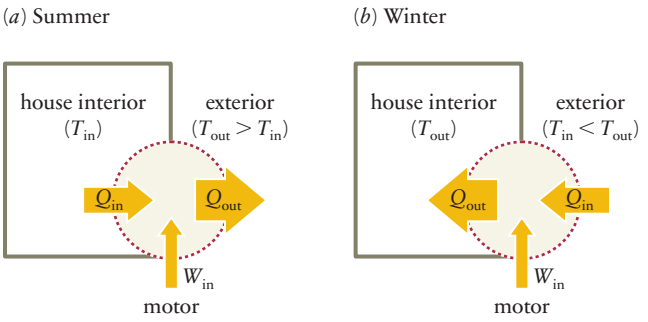


Figure 21.26 A heat pump can serve as an air-conditioner in summer by removing energy from inside the house and as a heater in winter by transferring energy thermally from outside to inside the house.



refrigerator is the reservoir at T_{out} (coils at the back or bottom of the refrigerator transfer the energy Q_{out} to the air).

Heat pumps are also used to heat and cool houses. In the summer, a heat pump operates as an air-conditioner and cools a house by moving energy from inside the house to outside (from cold to hot; **Figure 21.26*a***). In the winter, the same heat pump heats by moving energy from outside the house to inside (again from cold to hot, Figure 21.26*b*).

In designing a heat pump for cooling, we want to maximize the amount of energy transferred thermally from the low-temperature reservoir for a given amount of work done on the heat pump. The **coefficient of performance of cooling** of a heat pump is defined as the ratio of the thermal input of energy (the energy transferred thermally from the space that must be cooled) to the work done on the heat pump.

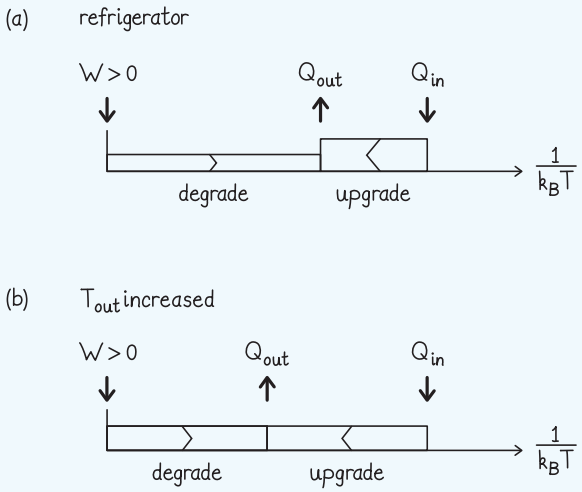
EXAMPLE 21.4 Cooling performance

For a set temperature T_{in} of a refrigerator interior, does the coefficient of performance of cooling increase, decrease, or stay the same as the temperature T_{out} of the room that contains the refrigerator increases? Assume the refrigerator is a reversible heat pump.

❶ **GETTING STARTED** I begin by drawing an entropy diagram for the refrigerator (**Figure 21.27*a***). Because the refrigerator is reversible, the areas of the upgrade and degrade rectangles are equal. The coefficient of performance of cooling is the ratio of the thermal input of energy Q_{in} (the height of the upgrade rectangle) to the work W done on the refrigerator (the height of the degrade rectangle).

❷ **DEVISE PLAN** To determine how the coefficient of performance of cooling changes as T_{out} increases, I must draw another entropy diagram, shifting the thermal output (the point where the two rectangles meet) to the left (Figure 21.27*b*). I must still have the area of the upgrade rectangle equal the area of the degrade rectangle, so in this diagram, I increase the height of the degrade rectangle and decrease that of the upgrade rectangle.

Figure 21.27



❸ **EXECUTE PLAN** The height of the upgrade rectangle corresponds to Q_{in}, and the height of the degrade rectangle corresponds to W. Shifting Q_{out} leftward to represent the increased T_{out} made the height of the upgrade rectangle (Q_{in}) smaller and the height of the degrade rectangle (W) greater. The ratio Q_{in}/W therefore decreases, which means that the coefficient of performance of cooling decreases. ✔

❹ **EVALUATE RESULT** The fact that the coefficient of performance of cooling decreases means that more energy is required to cool the refrigerator interior when the air outside the refrigerator is at a higher temperature, which makes sense intuitively. The reason is that when the refrigerator discards energy to the environment at a higher temperature, a given amount of energy causes a smaller entropy increase than when the energy is discarded at a lower temperature. This degradation of energy therefore compensates a smaller upgrade, reducing the coefficient of performance of cooling.

In designing a heat pump for heating, we want to maximize the thermal output of energy to the high-temperature reservoir for a given amount of work done on the heat pump. The **coefficient of performance of heating** of a heat pump is defined as the ratio of the thermal output of energy (the energy delivered thermally to the space that must be heated) to the work done on the heat pump.

Ordinary heating transfers energy from a high-temperature reservoir (say, burning fuel) to a low-temperature reservoir (a room to be heated, say), and the thermal input of energy Q_{in} is equal to the thermal output of energy Q_{out} (Figure 21.11). Equivalently, in Figure 21.14, $W = W_{in} > 0$ is converted to an identical amount of energy Q_{out}. In a heat pump, however, $Q_{out} > Q_{in}$, as Figure 21.26 indicates, because $Q_{out} = Q_{in} + W$. Note that while heat engines have an efficiency less than 1, the coefficient of performance of heating always exceeds 1. A typical heat pump has a coefficient of performance of heating of 3 to 4.

CONCEPTS

21.11 Estimate these quantities for the heat pump in Figure 21.26: (*a*) coefficient of performance of cooling and (*b*) coefficient of performance of heating.

21.4 Thermodynamic cycles

Heat engines and heat pumps are made up of three important components: a collection of moving parts, thermal reservoirs, and a *working substance* that undergoes a series of thermodynamic processes. For example, in the (heat) engine of an automobile, a mixture of air and gasoline is the working substance that produces the motion of the pistons that propels the automobile. Thermal energy generated in the combustion is transferred thermally to the environment through the engine's exhaust, and so the environment serves as a thermal reservoir. In a power plant, the working substance is water, which drives the blades of steam turbines. The steam boiler and the outside air serve as thermal reservoirs that transfer energy thermally to and from the water.

In this section we look at the thermodynamic processes that the working substance undergoes during the operation of a heat engine or heat pump. As we have seen, the engine or pump is a steady device: It must periodically return to a certain initial state. The working substance must therefore undergo a *cyclic process*—a sequence of steps in which the working substance begins and ends in the same thermal equilibrium state. Under those circumstances, the energy and the entropy of the working substance return to their initial values in one cycle.

Figure 21.28*a* illustrates a cyclic process that involves an expansion from state 1 to state 2 and then a compression back to state 1 via another path. We saw in Chapter 20 that during an expansion the work done on the system is negative and during a compression the work done on the system is positive. If the expansion takes place at a lower pressure than the compression (as it does in Figure 21.28*a*), the work done on the system during the expansion is smaller in magnitude than the work done during the compression. Therefore the work done on the system during a complete cycle from 1 to 2 and back to 1 is positive.

As we saw in Chapter 20, the magnitude of the negative work done on the system during the expansion $1 \rightarrow 2$ is equal to the hatched area in Figure 21.28*b*. The magnitude of the positive work done on the system during the compression $2 \rightarrow 1$ is equal to the shaded area in Figure 21.28*c*. Adding these two areas and taking their signs into account, we see that the work done on the system during a complete cycle is equal to the shaded area enclosed by the path that represents the cycle, as illustrated in 21.28*d*.

21.12 If you begin at state 1 again but this time go through the cycle in Figure 21.28 clockwise rather than counterclockwise, what are the magnitude and sign of (*a*) the work done on the system and (*b*) the work done on the environment?

Figure 21.28 and Checkpoint 21.12 demonstrate an important point:

The magnitude of the work done on a system during a cycle is equal to the area enclosed by the path that represents the cycle in a *PV* diagram. The work done on the system is negative for a clockwise cycle and positive for a counterclockwise cycle.

Because heat engines do positive work on the environment, their cycles are clockwise. Conversely, the environment

CONCEPTS

Figure 21.28 The work done on a system during a thermodynamic cyclic process is equal in magnitude to the area enclosed by the pathways making up the cycle.

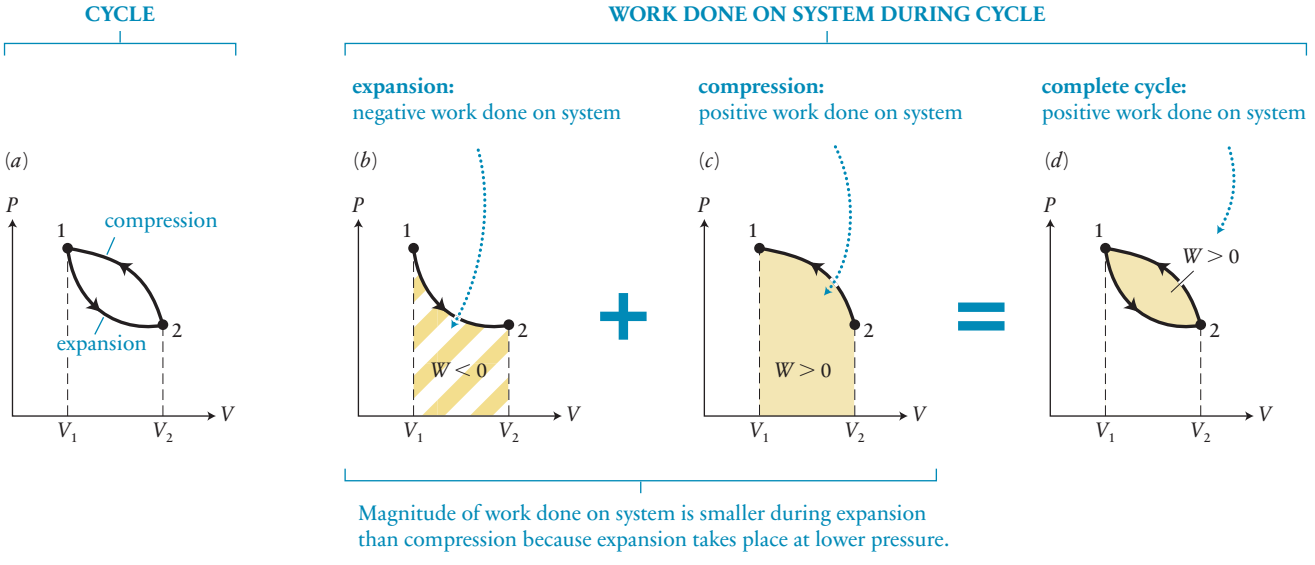
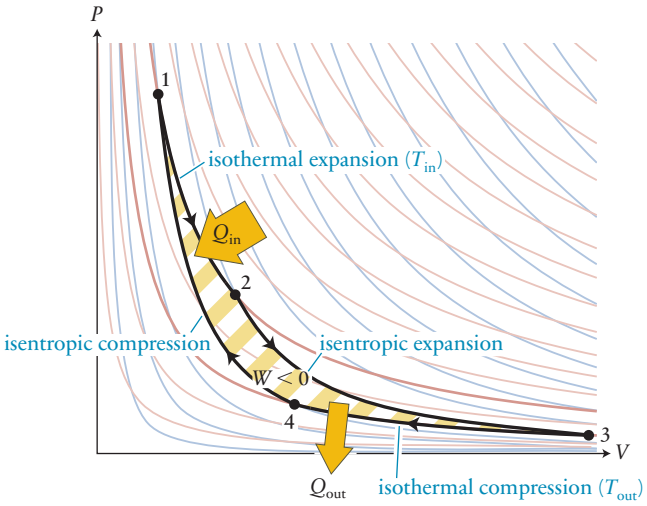


Figure 21.29 *PV* diagram for the thermodynamic process known as the Carnot cycle.



does positive work on heat pumps, and so their cycles are counterclockwise.

What thermodynamic processes occur in a heat engine? To answer this question, let us consider first the thermal input of energy from the high-temperature reservoir. This process takes place at a constant temperature T_{in} and so is isothermal. The thermal output of energy at temperature $T_{out} < T_{in}$ is also isothermal. It is not possible to construct a cycle from two isotherms, so we need other processes to connect the two isotherms. Because no other thermal exchange of energy takes place, these other processes we are looking for must be isentropic. **Figure 21.29** shows a thermodynamic cycle that consists of two isothermal processes ($1 \rightarrow 2$ and $3 \rightarrow 4$) and two isentropic processes ($2 \rightarrow 3$ and $4 \rightarrow 1$). Because of its importance in thermodynamics, this cycle is called the **Carnot cycle** after the French physicist Sadi Carnot (1796–1832), who produced the first successful description of the operation of heat engines.

Beginning in state 1 in Figure 21.29, the working substance expands isothermally at a temperature T_{in} until it reaches some state 2 (at which point its volume has increased). The working substance then undergoes an isentropic expansion during which its temperature drops from T_{in} to T_{out}. From this state 3, the working substance is compressed isothermally to state 4, from where it undergoes an isentropic compression back to state 1, completing the cycle.

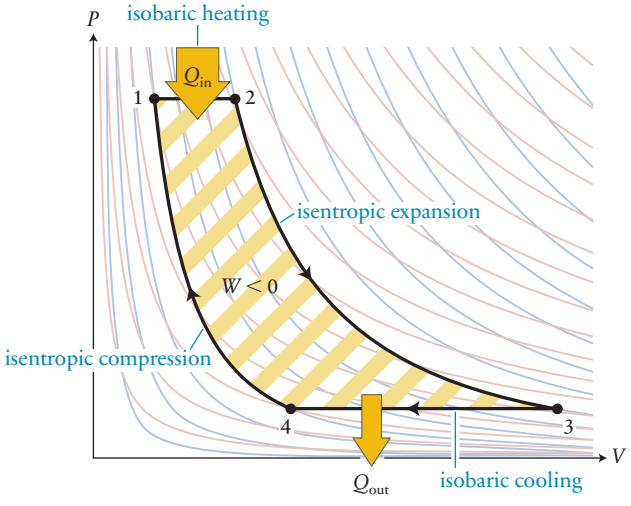
Energy is transferred thermally into and out of the system only during the two isothermal processes, but all four processes involve work. The work done on the system during the cycle is equal to the area enclosed by the cycle and has been shaded in Figure 21.29. Because W is negative, the shading is hatched. (Remember that is what we want: The heat engine must do positive work *on its environment*, and so the work done *on the system* is negative.)

The Carnot cycle has a number of deficiencies that makes it impractical for a heat engine. First of all, as you can see from the small hatched area in Figure 21.29, the amount of work done on the environment in each cycle is small. More important, the presence of the isotherms means that the cycle cannot be repeated quickly. As we saw in Section 20.7, isothermal processes are very slow, and so an engine that operates on the Carnot cycle would put out mechanical energy at a very low rate; that is, its power would be very small.

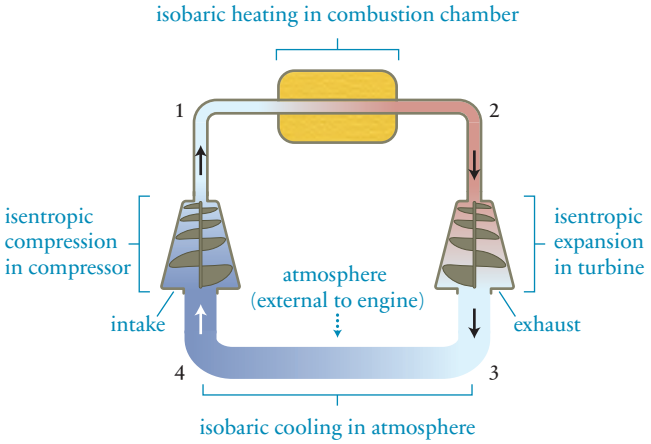
To alleviate these problems, we want to avoid isotherms and "open up" the cycle so that more work is done on the environment. If we replace the two isotherms in Figure 21.29 by isobars, we obtain the cycle shown in **Figure 21.30*a***. This cycle, called the **Brayton cycle**, does substantially more work on the environment than the Carnot cycle and can be carried

Figure 21.30 The thermodynamic process known as the Brayton cycle.

(*a*) *PV* diagram for Brayton cycle



(*b*) Flow chart for gas turbine operating on Brayton cycle



CONCEPTS

out fairly quickly. Gas turbines and jet engines operate on a cycle that is closely approximated by the Brayton cycle.

In the Brayton cycle, the working substance undergoes two isobaric processes ($1 \rightarrow 2$ and $3 \rightarrow 4$) and two isentropic processes ($2 \rightarrow 3$ and $4 \rightarrow 1$). Energy thermally enters and leaves the system during the isobaric processes. If we begin and end the isentropic compression in Figures 21.29 and 21.30 from the same states 4 and 1, the hatched area enclosed by the Brayton cycle is much larger than the area enclosed by the Carnot cycle. The larger hatched area means that a steady device that operates on the Brayton cycle has a much greater mechanical output of energy than a steady device that operates on the Carnot cycle. Also, the isobaric and isentropic processes can be carried out rapidly, which means the power output in a Brayton cycle can be high.

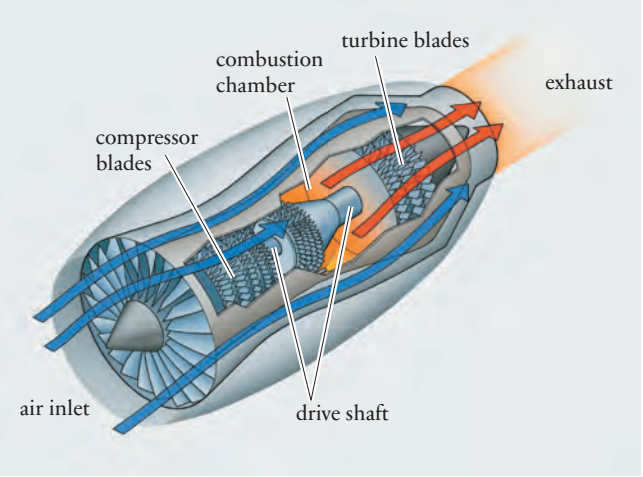
A gas turbine engine that operates on the Brayton cycle has three main components: a gas compressor, a combustion chamber, and an expansion turbine. In practice, such an engine is an open system: Fresh air flows in at one end, and hot gases are expelled at the other end. For the purpose of thermodynamic analysis, however, we can pretend that the exhaust gases are reused in the intake.

Figure 21.30*b* shows a flow chart for a gas turbine engine. Gas—typically ambient air—is drawn into the compressor, which isentropically compresses the air to high pressure $4 \rightarrow 1$. The compressed air then flows into the combustion chamber, where it is heated to temperatures as high as 2000 K at constant pressure ($1 \rightarrow 2$). The heated, pressurized air then expands isentropically ($2 \rightarrow 3$), and the hot gases are expelled at high speed from the rear of the engine, driving the engine forward. The exhaust gases then cool back down to the intake temperature—another constant-pressure process ($3 \rightarrow 4$)—and the cycle is complete.

To power the compressor, a turbine is placed in the expansion chamber. The expanding gases drive this turbine, which in turn feeds some energy from the expanding gases to the compressor. In a jet engine, for example, the turbine and the compressor are placed on a single drive shaft, so the turbine directly feeds the compressor (**Figure 21.31**).

Although the Brayton cycle delivers more energy and power than the Carnot cycle, it is not as efficient, which means that the work done on the environment divided by the thermal input of energy is smaller. The reason for this diminished efficiency is illustrated in **Figure 21.32*a***. In the Brayton cycle, the thermal input and output of energy do not take place at fixed temperatures as in the Carnot cycle. The Brayton thermal input of energy starts at the temperature T_1 reached during compression and then moves up to a higher

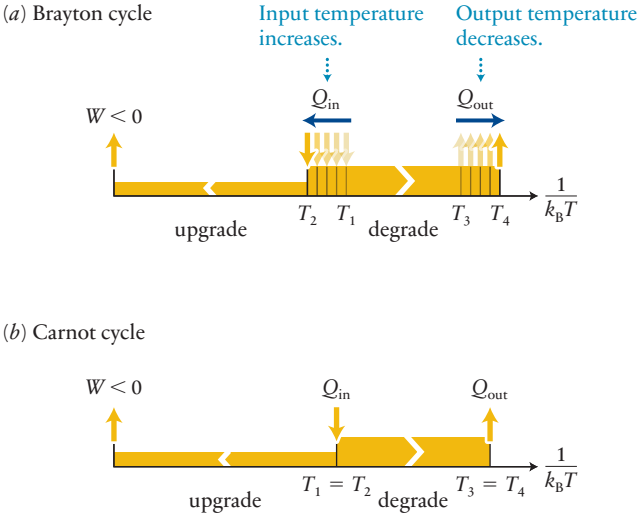
Figure 21.31 Jet engine.



temperature T_2 during combustion. The thermal output of energy starts at temperature T_3 (the temperature at the end of the isentropic expansion) and then slides down to the lower temperature T_4 (the temperature at the beginning of the isentropic compression).

As we saw in Section 21.3, the efficiency of a heat engine is highest when the thermal input of energy takes place at the highest possible temperature and the output at the lowest possible temperature. So the part of the cycle where thermal input of energy occurs at T_2 and thermal output of energy occurs at T_4 has the greatest efficiency. All other parts of the cycle have lower efficiency, which contributes to an overall efficiency that is lower than the efficiency we would see if *all* energy exchange had taken place at only the highest and lowest temperatures, T_2 and T_4, respectively. By contrast, thermal input of energy at one

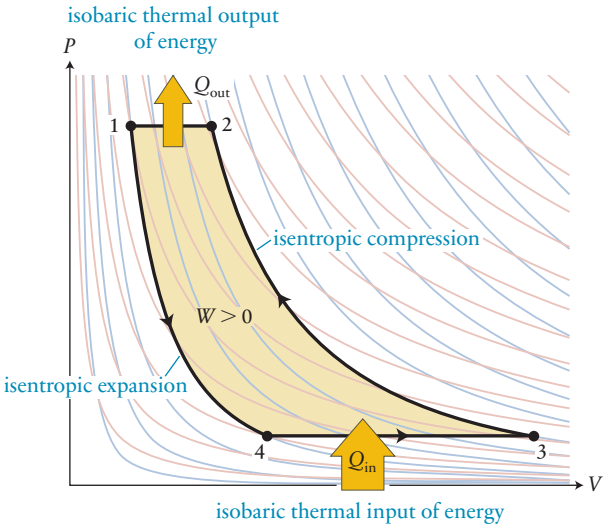
Figure 21.32 Entropy diagrams for Brayton and Carnot cycles.



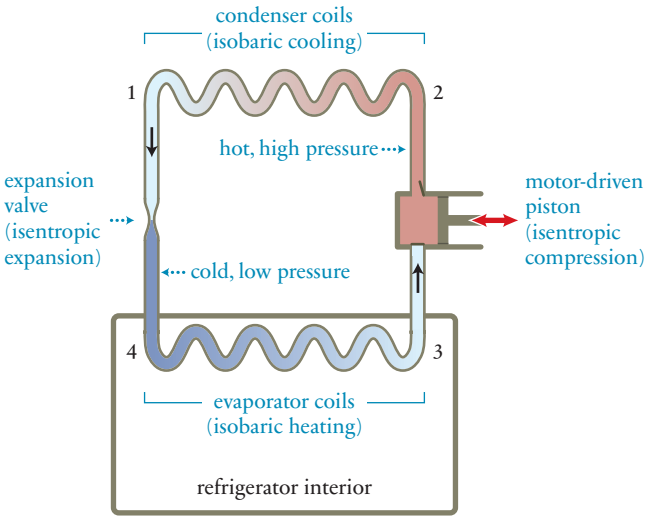
CONCEPTS

Figure 21.33 A Brayton cycle run in reverse is the cycle for a heat pump.

(*a*) *PV* diagram for Brayton cycle run in reverse



(*b*) Flow chart for refrigerator operating on reverse Brayton cycle



fixed high temperature and output at one fixed low temperature is exactly what happens in a Carnot cycle (Figure 21.32*b*).

By running the Brayton cycle in reverse (**Figure 21.33*a***), we obtain a heat pump. The cycle used in refrigerators, air-conditioners, and home heat pumps is a bit more complex because the working substance in these devices changes from liquid to vapor and back. However, if we assume the working substance is a gas, the basic thermodynamic processes are the isentropic ($3 \rightarrow 2$ and $1 \rightarrow 4$) and isobaric ($2 \rightarrow 1$ and $4 \rightarrow 3$) ones illustrated in Figure 21.33*a*.

Figure 21.33*b* shows a flow chart of the reverse Brayton cycle that takes place in a refrigerator. The working substance is first isentropically compressed by a motor-driven piston ($3 \rightarrow 2$). The compression raises both the temperature and the pressure in the working substance, which is then cooled down at constant pressure in a long coiled tube called the *condenser* ($2 \rightarrow 1$). This tube is exposed to the ambient temperature. After passing through the condenser, the high-pressure and now somewhat cooler working substance passes through an expansion valve ($1 \rightarrow 4$) and expands isentropically. In the process of expanding, the working substance cools significantly. It then passes through a second set of coils called the *evaporator* ($4 \rightarrow 3$). Because the temperature of the space around the evaporator is higher than the temperature of the working substance in the evaporator, energy is thermally transferred from the surrounding space to the evaporator. As Figure 21.33*b* shows, this space surrounding the evaporator is the interior of the refrigerator, where food is stored and thus kept cold. After passing through the evaporator, the working substance is returned to the compressor, where the cycle repeats.

Note that a heat pump transfers energy from the interior of the refrigerator at low temperature to the room at a higher temperature. To accomplish this, it needs mechanical energy to first raise the temperature of the working substance above that of the room so that energy can be transferred thermally out of the hot working substance into the room. Then the temperature of the working substance is lowered to below that of the interior of the refrigerator so that energy can be transferred thermally from the interior of the refrigerator into the working substance.

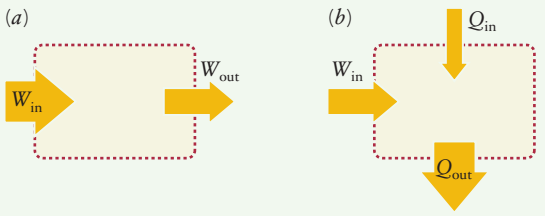
21.13 Why doesn't the transfer of energy from the interior of the refrigerator to the room violate the entropy law?

CONCEPTS

Self-quiz

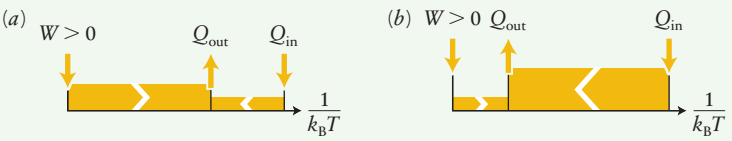
1. Describe the energy transfers in **Figure 21.34**. What is implied about the energy of the system?

Figure 21.34



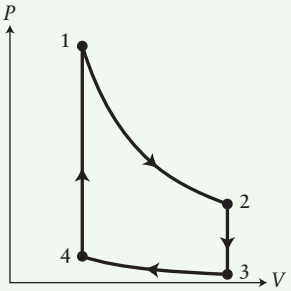
2. Describe the energy conversions in **Figure 21.35**. Which are possible?

Figure 21.35



3. Write the ratios for the efficiency of a heat engine, the coefficients of performance of cooling for a refrigerator, and the coefficients of performance of heating a room using a heat pump in terms of W, Q_{in}, and Q_{out}. Make a generalization about the quantity on the top of each ratio and the quantity on the bottom of each ratio.
4. **Figure 21.36** shows an ideal gas cycle consisting of two isotherms and two isochores. Complete the table by writing the algebraic signs of each of the terms ΔE, W, and Q. If a term is zero, write zero.

Figure 21.36



	ΔE	W	Q
$1 \rightarrow 2$			
$2 \rightarrow 3$			
$3 \rightarrow 4$			
$4 \rightarrow 1$			

Answers

1. (*a*) More work is done on the system than the system does on the environment, so the energy of the system increases. (*b*) More energy is transferred thermally out of the system than into it, and work is done on the system. Depending on how the work done on the system compares to the difference between the thermal transfers of energy, the energy of the system can increase, decrease, or stay the same.
2. In both cases mechanical energy is degraded into thermal energy of a high-temperature reservoir while thermal energy of a low-temperature reservoir is upgraded to thermal energy of a high-temperature reservoir. (*a*) Because the area of the degrade rectangle is larger than the area of the upgrade rectangle, entropy is increasing, so this process is possible. (*b*) Because the area of the upgrade rectangle is larger than the area of the degrade rectangle, entropy is decreasing, so this process is not possible.
3. Efficiency $= -W/Q_{in}$, the coefficient of performance of cooling $= Q_{in}/W$, and the coefficient of performance of heating $= Q_{out}/W$. The quantities in the numerator are the desired results: work done by the engine on the environment, energy transferred thermally out of the contents of the refrigerator, and energy transferred thermally into the room to be heated. The quantities in the denominator are the energies required to obtain the desired result: the thermal transfer of energy required to run the engine and the work required to run the refrigerator or heat pump.
4. ΔE: $1 \rightarrow 2 = 0$ (isotherm); $2 \rightarrow 3 = -$ (temperature decreases); $3 \rightarrow 4 = 0$ (isotherm); $4 \rightarrow 1 = +$ (temperature increases). W: $1 \rightarrow 2 = -$ (expansion); $2 \rightarrow 3 = 0$ (isochoric process); $3 \rightarrow 4 = +$ (compression); $4 \rightarrow 1 = 0$ (isochoric process). Q (sign follows from signs of ΔE and W): $1 \rightarrow 2 = +$; $2 \rightarrow 3 = -$; $3 \rightarrow 4 = -$; $4 \rightarrow 1 = +$.

21.5 Entropy constraints on energy transfers

In the first part of this chapter, we used entropy diagrams to describe the energy conversions that occur in heat engines and pumps. We stated, without proof, that the area of the rectangle for any energy conversion or transfer represented in such a diagram is a measure of the corresponding entropy change. In this section we develop the quantitative framework for computing entropy changes from entropy diagrams.

Steady devices return to the same state after one cycle, so that

$$\Delta E = 0 \quad \text{(steady device, complete cycle)}. \tag{21.1}$$

The energy law (Eq. 20.1) then tells us that for a steady device $Q + W = 0$. Because $Q = Q_{\text{in}} - Q_{\text{out}}$, we see that over one cycle the work done on a steady device must be equal to the difference in the thermal output and input of energy:

$$W = Q_{\text{out}} - Q_{\text{in}} \quad \text{(steady device, complete cycle)}. \tag{21.2}$$

Because the initial and final states are the same, the entropy of a steady device doesn't change either:

$$\Delta S_{\text{dev}} = 0 \quad \text{(steady device, complete cycle)}. \tag{21.3}$$

From Chapter 19 we know that only processes in which the entropy either increases or remains unchanged can occur:

$$\Delta S_{\text{dev}} + \Delta S_{\text{env}} \geq 0. \tag{21.4}$$

Equation 21.3 thus means that we must have

$$\Delta S_{\text{env}} \geq 0 \quad \text{(steady device, complete cycle)}. \tag{21.5}$$

As we saw in the first part of this chapter, the entropy constraint in Equation 21.5 limits the type of energy transfers that can occur in a steady device.



21.14 What is the change in the energy of the environment in a device for which Eqs. 21.1–21.5 hold?

We begin by considering a steady device that transfers energy thermally from a reservoir at temperature T_{in} to a reservoir at a lower temperature $T_{\text{out}} < T_{\text{in}}$ (**Figure 21.37*a***).

In Chapter 20 we saw that the entropy change associated with energy transferred thermally at a fixed temperature is given by (Eq. 20.54)

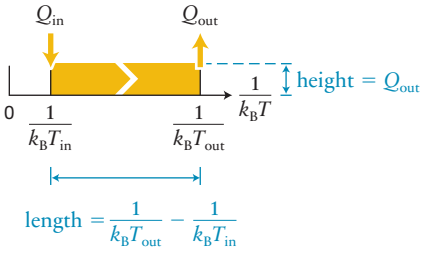
$$\Delta S = \frac{Q}{k_{\text{B}}T} \quad \text{(isothermal process)}. \tag{21.6}$$

The entropy diagram in Figure 21.37*a* shows two thermal transfers of energy—a transfer out of the environment at the high-temperature reservoir and a transfer into the environment at the low-temperature reservoir. Let us first apply Eq. 21.6 to the high-temperature reservoir. Because energy is transferred thermally out of that reservoir, $Q_{\text{H}} = -Q_{\text{in}} < 0$. Equation 21.6 thus yields a negative entropy change of the high-temperature reservoir, which is at temperature T_{in}:

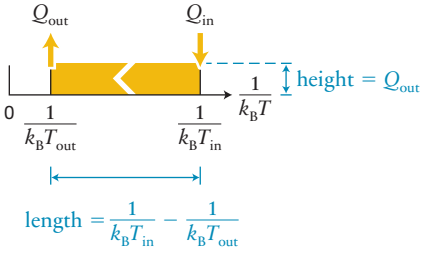
$$\Delta S_{\text{H}} = -\frac{Q_{\text{in}}}{k_{\text{B}}T_{\text{in}}}. \tag{21.7}$$

Figure 21.37 Entropy diagram for thermal transfer of energy (*a*) from a high-temperature reservoir to a low-temperature reservoir and (*b*) from a low-temperature reservoir to a high-temperature reservoir.

(*a*) Transfer to lower-temperature reservoir



(*b*) Transfer to higher-temperature reservoir



QUANTITATIVE TOOLS

Likewise, for the low-temperature reservoir, which is at temperature T_{out}, $Q_L = Q_{out} > 0$, so the entropy change is positive:

$$\Delta S_L = \frac{Q_{out}}{k_B T_{out}}. \tag{21.8}$$

Adding the entropy changes in the environment, we get for the environment entropy change during the complete cycle represented in the entropy diagram

$$\Delta S_{env} = \Delta S_H + \Delta S_L = -\frac{Q_{in}}{k_B T_{in}} + \frac{Q_{out}}{k_B T_{out}} \quad \text{(steady state)}. \tag{21.9}$$

Because $\Delta E = 0$ (Eq. 21.1) and because no work is involved, the energy law tells us that $\Delta E = Q_{in} - Q_{out} = 0$, so $Q_{in} = Q_{out}$. Substituting this result in Eq. 21.9, we obtain for the entropy change in the environment

$$\Delta S_{env} = -\frac{Q_{out}}{k_B T_{in}} + \frac{Q_{out}}{k_B T_{out}} = Q_{out}\left(\frac{1}{k_B T_{out}} - \frac{1}{k_B T_{in}}\right) \quad \text{(steady state)}. \tag{21.10}$$

The right side of this equation gives exactly what we have been calling the area of the rectangle representing the energy transfer in the entropy diagram. As Figure 21.37*a* shows, the height of this rectangle is Q_{out} (a positive quantity):

$$\text{height of rectangle} = Q_{out}. \tag{21.11}$$

If we use the quantity $1/k_B T$ to quantify the horizontal scale in the diagram, we see that the left edge of the rectangle is at $1/k_B T_{in}$ and the right edge is at $1/k_B T_{out}$. The length of the rectangle is thus

$$\text{length of rectangle} = \frac{1}{k_B T_{out}} - \frac{1}{k_B T_{in}}. \tag{21.12}$$

Because $T_{out} < T_{in}$, this quantity is positive. Note that the length is exactly equal to the term in parentheses in Eq. 21.10, so the product of the height and the length—the area—of the rectangle indeed represents the entropy change of the environment ΔS_{env} due to the transfer of energy from the high-temperature reservoir to the low-temperature reservoir.

Let us next consider the reverse process: a thermal transfer of energy from a low-temperature reservoir to a high-temperature reservoir. As we saw in Section 21.2, such an upgrade of energy is not possible. The entropy diagram for this impossible process is shown in Figure 21.37*b*. Note that the rectangle has the same length and height as the one in Figure 21.37*a*. It points to the left (from low temperature to high temperature), indicating that the entropy change is negative. Therefore we take the negative of the area of the rectangle to compute the entropy change in the environment:

$$\Delta S_{env} = -Q_{out}\left(\frac{1}{k_B T_{in}} - \frac{1}{k_B T_{out}}\right) \quad \text{(steady state)}. \tag{21.13}$$

Because the entropy change in the environment must be positive (Eq. 21.5), we see that this type of transfer cannot occur by itself.

The procedure for computing the entropy change from an entropy diagram is presented in the Procedure box "Computing entropy changes from entropy diagrams."

QUANTITATIVE TOOLS

PROCEDURE: Computing entropy changes from entropy diagrams

To determine the entropy change in the environment from an entropy diagram, compute the area of the rectangles that represent the energy transfers as follows:

1. Determine the length of the horizontal part of each rectangle by subtracting the position of the left end of the rectangle on the $1/k_BT$ scale from the position of the right end on the scale. The position of mechanical energy is always at the zero point on this scale. Verify that the result is a positive quantity. (If it isn't, you made a mistake somewhere.)
2. Express the height of each rectangle in terms of Q_{in}, Q_{out}, and W. Remember that the height must be a positive quantity; Q_{in} and Q_{out} are always positive, but W can be positive or negative. Avoid using the input or output where two rectangles meet; if necessary, you can use Eq. 21.2, $W = Q_{out} - Q_{in}$, to express the height in terms of the quantities that are given in the problem.
3. Determine the sign of the entropy change. If the rectangle points left (energy upgrade), the entropy change is negative. If the rectangle points right (energy degradation), the entropy change is positive.
4. The entropy change for each rectangle is given by the area of the rectangle (the length multiplied by the height computed in steps 1 and 2) preceded by the algebraic sign determined in step 3:

$$\Delta S = (\text{sign}) \times (\text{length}) \times (\text{width}).$$

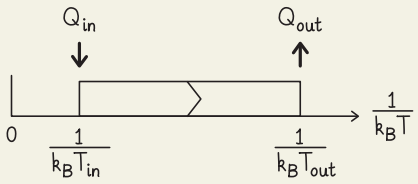
5. The entropy change for the process is the sum of the individual entropy changes you have calculated.

EXAMPLE 21.5 Thermal transfer of energy

A steel bar transfers energy from a thermal reservoir at 750 K to one at 520 K. If the bar transfers 1.5×10^6 J of energy during a certain time interval over which the temperature distribution of the bar does not change, what is the entropy change of the environment?

❶ **GETTING STARTED** Because the temperature distribution does not change, the state of the bar remains constant and I can consider the bar to be a steady device. I begin by drawing an entropy diagram (**Figure 21.38**), showing a right-pointing arrow in the rectangle because the direction of the energy transfer is from the higher-temperature reservoir to the lower-temperature one. I am given $T_{in} = 750$ K, $T_{out} = 520$ K, and the amount of energy transferred $Q_{in} = Q_{out} = 1.5 \times 10^6$ J.

Figure 21.38



❷ **DEVISE PLAN** To determine the entropy change of the environment, I follow the procedure shown above that tells me how to compute the entropy change from my entropy diagram.

❸ **EXECUTE PLAN** Because the rectangle arrow points to the right, I know that the entropy of the environment increases and thus the entropy change is positive. The height of the rectangle is Q_{in}. The length is $[1/k_BT_{out} - 1/k_BT_{in}]$. The entropy change is thus

$$\Delta S_{env} = +Q_{in}\left[\frac{1}{k_BT_{out}} - \frac{1}{k_BT_{in}}\right] = \frac{+Q_{in}}{k_B}\left[\frac{1}{T_{out}} - \frac{1}{T_{in}}\right].$$

Substituting the values given, I obtain

$$\Delta S_{env} = \frac{1.5 \times 10^6 \text{ J}}{1.38 \times 10^{-23} \text{ J/K}}\left[\frac{1}{520 \text{ K}} - \frac{1}{750 \text{ K}}\right]$$

$$= 6.4 \times 10^{25}. \checkmark$$

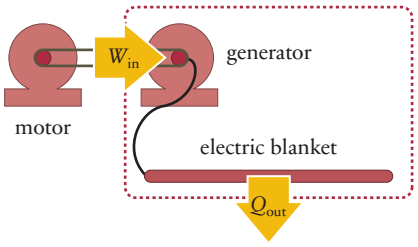
❹ **EVALUATE RESULT** Energy is degraded (hot to cold), so I expect a positive entropy change. The number I obtained for the magnitude of the entropy change is very large, but I know from Chapter 19 that entropy values are very large, and so my answer is not unreasonable.

21.15 Suppose the bar in Example 21.5 is disconnected from the high-temperature reservoir and allowed to thermally equilibrate with the low-temperature reservoir. (*a*) Are the changes in the energy and entropy of the bar during this process positive, negative, or zero? (*b*) Is there any change in the entropy of the environment? (*c*) Does the sum $S_{bar} + S_{env}$ change?

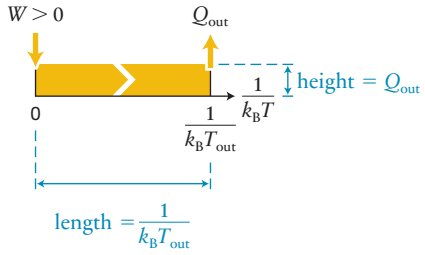
Next we consider a device that converts mechanical energy (work) to thermal energy. In Section 21.1 we discussed a paddle wheel immersed in water. Now we consider a somewhat different situation: a motor doing work on an electric generator that delivers electrical energy to an electric blanket,

Figure 21.39 A setup that converts mechanical energy (from the motor) to thermal energy.

(*a*)



(*b*)



which converts the electrical energy to thermal energy (**Figure 21.39*a***). We define the generator, blanket, and wires that connect the two as our system. The motor (which is part of the environment) does work on the system, and energy is transferred thermally by the blanket (part of the system) to the environment. We assume that the temperatures of the blanket and the environment remain constant. This assumption ensures that the initial and final states of the system are the same.

Figure 21.39*b* shows the entropy diagram for this process. Because the transfer is to the right, the entropy change is positive. The height of the rectangle is Q_{out}. (Because $W = Q_{out}$ we could also write the height of the rectangle as W.) The rectangle stretches from the origin (0 on the $1/k_B T$ scale) to $1/k_B T_{out}$, so its length is $1/k_B T_{out}$. The entropy change is thus

$$\Delta S_{env} = +Q_{out}\frac{1}{k_B T_{out}} = \frac{Q_{out}}{k_B T_{out}} \quad \text{(steady state)}. \qquad (21.14)$$

This is indeed the result we would expect: No entropy change is associated with the work done on the system, and the entropy change caused by the thermal transfer of energy to the environment is $Q_{out}/k_B T_{out}$, which is exactly the result we obtain in Eq. 21.14.

21.16 (*a*) Draw an entropy diagram for a steady device that converts thermal energy from a thermal reservoir to mechanical energy. (*b*) From the diagram, compute the entropy change that occurs during the process and show that such a device is impossible.

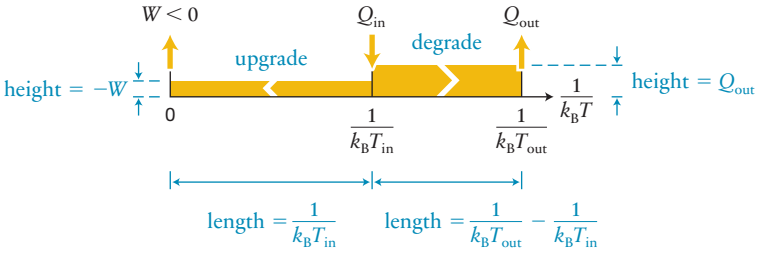
21.6 Heat engine performance

Consider a heat engine that takes in energy thermally from a high-temperature reservoir, converts part of this energy to mechanical energy by doing work on the environment ($W < 0$), and discards the remainder to a low-temperature reservoir. As we saw in Section 21.3, the entropy diagram has two rectangles: one for the energy upgrade and the other for the degradation (**Figure 21.40**). To determine the entropy change in the environment over one complete cycle, we add the entropy changes of the two rectangles:

$$\Delta S_{env} = \Delta S_{upgrade} + \Delta S_{degrade} \quad \text{(complete cycle)}. \qquad (21.15)$$

The entropy change for the degradation is given by Eq. 21.10, so we need only calculate the change for the upgrade. The height of the upgrade rectangle is $-W$. Because $W = Q_{out} - Q_{in}$ (Eq. 21.2), the height of the rectangle can also be written as $Q_{in} - Q_{out}$. The rectangle stretches from the origin (0 on the $1/k_B T$ scale) to $1/k_B T_{in}$, so its length is $1/k_B T_{in}$. Because

Figure 21.40 Entropy diagram for a heat engine that upgrades a portion of the thermal energy to mechanical energy and degrades the rest.



QUANTITATIVE TOOLS

the transfer is to the left, the entropy change is negative. The entropy change associated with the upgrade is thus

$$\Delta S_{\text{upgrade}} = -(Q_{\text{in}} - Q_{\text{out}})\left[\frac{1}{k_B T_{\text{in}}}\right]. \tag{21.16}$$

Substituting Eqs. 21.16 and 21.10 into Eq. 21.15 gives

$$\Delta S_{\text{env}} = -(Q_{\text{in}} - Q_{\text{out}})\left[\frac{1}{k_B T_{\text{in}}}\right] + Q_{\text{out}}\left[\frac{1}{k_B T_{\text{out}}} - \frac{1}{k_B T_{\text{in}}}\right], \tag{21.17}$$

or

$$\Delta S_{\text{env}} = -\frac{Q_{\text{in}}}{k_B T_{\text{in}}} + \frac{Q_{\text{out}}}{k_B T_{\text{out}}} \quad \text{(complete cycle)}, \tag{21.18}$$

which is exactly Eq. 21.9, the expression we derived for the entropy change for a simple device that thermally transfers energy from a high-temperature reservoir to a low-temperature reservoir. The reason we obtain the same expression for the more complicated case of a heat engine is that no entropy change is associated with mechanical transfers of energy. Only the thermal transfers of energy to and from the thermal reservoirs contribute to the entropy change.

For a reversible heat engine, the entropy change in the environment must be equal to zero over one cycle (everything is back to what it was before), $\Delta S_{\text{env}} = 0$, so

$$-\frac{Q_{\text{in}}}{k_B T_{\text{in}}} + \frac{Q_{\text{out}}}{k_B T_{\text{out}}} = 0 \quad \text{(reversible heat engine)}. \tag{21.19}$$

Rearranging this equation, we get

$$\frac{Q_{\text{out}}}{Q_{\text{in}}} = \frac{T_{\text{out}}}{T_{\text{in}}} \quad \text{(reversible heat engine)}. \tag{21.20}$$

For a reversible heat engine that transfers energy from a thermal reservoir at T_{in}, the amount of energy discarded is proportional to T_{out}. The amount of energy discarded is independent of the type of working substance used and depends on only the temperatures of the two reservoirs. If we lower T_{out}, we decrease the amount of discarded energy, which means a greater portion of the input energy can be converted to mechanical energy. So we can maximize the work done on the environment by minimizing the temperature of the low-temperature reservoir.

21.17 Is it possible to reduce the amount of thermally discarded energy to zero by bringing T_{out} down to absolute zero?

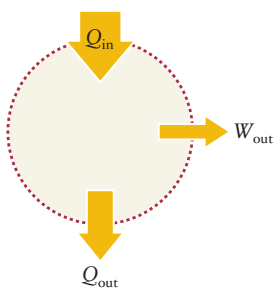
Figure 21.41 shows the energy input-output diagram for a heat engine. In Section 21.3 we defined the efficiency of such an engine as the ratio of the work done on the environment ($-W$) to the thermal input of energy (Q_{in}). Using the symbol η (lowercase Greek eta) for efficiency, we can write

$$\eta \equiv \frac{-W}{Q_{\text{in}}}. \tag{21.21}$$

Substituting W from Eq. 21.2 yields

$$\eta = \frac{Q_{\text{in}} - Q_{\text{out}}}{Q_{\text{in}}} = 1 - \frac{Q_{\text{out}}}{Q_{\text{in}}}. \tag{21.22}$$

Figure 21.41 Energy input-output diagram for a heat engine.



Equation 21.22 is a consequence of the energy law and holds for both reversible and irreversible heat engines that involve transfers of energy to and from two thermal reservoirs.

Figure 21.23 shows that for a given amount of work $-W$ done by a heat engine, Q_{in} is smallest when the engine is reversible. Therefore the efficiency of a heat engine is maximum when it is reversible. For a reversible heat engine, we can substitute Eq. 21.20 into Eq. 21.22 to obtain an expression for the maximum efficiency:

$$\eta_{max} = 1 - \frac{T_{out}}{T_{in}} = \frac{T_{in} - T_{out}}{T_{in}} \quad \text{(reversible heat engine)}. \quad (21.23)$$

In practice, the efficiency of a heat engine is always smaller than this maximum efficiency because it is impossible to completely eliminate irreversible processes. The important point to remember, however, is that even for a reversible heat engine, the efficiency is less than 1. The maximum efficiency is obtained with a minimum transfer of thermal energy to the low-temperature reservoir and has *nothing* to do with the details of the workings of the heat engine. *Only* the temperatures of the two reservoirs have any effect on the maximum efficiency: It is greatest when the ratio T_{out}/T_{in} is smallest.

EXAMPLE 21.6 Environmental impact of power plant

A power plant burns fossil fuel to produce steam at 650 K. The pressurized steam drives a steam turbine, where the steam condenses to water and is discarded at 310 K. What is the minimum quantity of energy discarded thermally by a 1.0-GW power plant to the environment during one day?

❶ **GETTING STARTED** Because I am interested in the minimum quantity of thermal energy discarded, I can treat the turbine as a reversible heat engine. Energy is thermally transferred to the turbine from the steam, which serves as the high-temperature reservoir at $T_{in} = 650$ K. Energy is thermally discarded along with the condensed steam at $T_{out} = 310$ K.

❷ **DEVISE PLAN** To calculate the minimum quantity of energy discarded to the environment, $Q_{out,min}$, I can use Eq. 21.2, which tells me that for a given value of W, Q_{out} is minimum when Q_{in} is minimum. No value is given for Q_{in}, but Eq. 21.21 tells me that Q_{in} is minimum when η has its maximum value, which I can calculate using Eq. 21.23. In one day, a 1.0-GW power plant delivers an amount of energy equal to $(24\text{ h})(3600\text{ s}/1\text{ h})\,(1.0 \times 10^9\text{ J/s}) = 8.6 \times 10^{13}$ J, and so $W = -8.6 \times 10^{13}$ J.

❸ **EXECUTE PLAN** The maximum efficiency is (Eq. 21.23)

$$\eta_{max} = \frac{650\text{ K} - 310\text{ K}}{650\text{ K}} = 0.523,$$

and so the minimum quantity of energy transferred thermally to the turbine is

$$Q_{in,min} = \frac{-W}{\eta_{max}} = \frac{8.6 \times 10^{13}\text{ J}}{0.523} = 1.6 \times 10^{14}\text{ J}.$$

From Eq. 21.2 I thus calculate the minimum quantity of energy discarded thermally:

$$Q_{out,min} = Q_{in,min} + W = (1.6 \times 10^{14}\text{ J}) - (8.6 \times 10^{13}\text{ J})$$
$$= 7.8 \times 10^{13}\text{ J}. \checkmark$$

❹ **EVALUATE RESULT** The quantity of thermal energy leaving the turbine is smaller than the quantity of thermal energy entering it (as it should be) and approximately equal to the quantity of energy delivered by the plant. This makes sense because the maximum efficiency is close to $\frac{1}{2}$. I assumed the turbine to be reversible, which is unlikely. However, if the turbine is not reversible, more energy is dissipated, and so my result for the minimum amount of energy discarded still holds.

QUANTITATIVE TOOLS

21.18 A hurricane thermally takes in energy from ocean water at 26 °C, discards some of it to the upper atmosphere at −52 °C, and converts the remainder to mechanical energy in the form of wind. (*a*) What is the maximum possible efficiency for the thermal-to-mechanical conversion? (*b*) If the temperature of the water increases by 1 °C, by what amount does the maximum efficiency increase?

If we run a heat engine in reverse, we obtain a heat pump (Figure 21.42). When the heat pump is used as a heater, we define the **coefficient of performance of heating** as

$$\text{COP}_{\text{heating}} \equiv \frac{Q_{\text{out}}}{W}. \tag{21.24}$$

Because $W = W_{\text{in}} > 0$ and $W = Q_{\text{out}} - Q_{\text{in}}$, Q_{out} is always greater than or equal to W, and so $\text{COP}_{\text{heating}}$ for a heat pump is always greater than or equal to 1.

Substituting Eq. 21.2 into Eq. 21.24 gives

$$\text{COP}_{\text{heating}} = \frac{Q_{\text{out}}}{Q_{\text{out}} - Q_{\text{in}}} = \frac{1}{1 - Q_{\text{in}}/Q_{\text{out}}}. \tag{21.25}$$

For a reversible heat pump that transfers energy from a lower-temperature thermal reservoir at T_{in} to a higher-temperature reservoir at T_{out}, we can substitute Eq. 21.20 into Eq. 21.25 to obtain the maximum coefficient of performance of heating:

$$\text{COP}_{\text{heating,max}} = \frac{1}{1 - T_{\text{in}}/T_{\text{out}}} = \frac{T_{\text{out}}}{T_{\text{out}} - T_{\text{in}}} \quad \text{(reversible heat pump)}. \tag{21.26}$$

This expression tells us that the maximum coefficient of performance of heating is greatest when the ratio $T_{\text{in}}/T_{\text{out}}$ is close to 1.



21.19 What is the coefficient of performance of heating for the electric blanket in Figure 21.39?

When a heat pump is used to cool (as a refrigerator or air conditioner), we define the **coefficient of performance of cooling** as

$$\text{COP}_{\text{cooling}} \equiv \frac{Q_{\text{in}}}{W}. \tag{21.27}$$

Substituting Q_{in} from Eq. 21.2 gives

$$\text{COP}_{\text{cooling}} = \frac{Q_{\text{out}} - W}{W} = \frac{Q_{\text{out}}}{W} - 1 = \text{COP}_{\text{heating}} - 1. \tag{21.28}$$

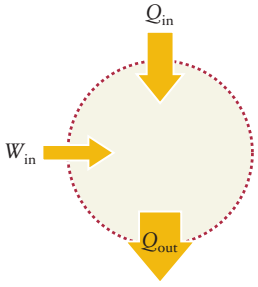
For a reversible heat pump, we can substitute Eq. 21.26 into Eq. 21.28 to obtain an expression for the maximum coefficient of performance of cooling:

$$\text{COP}_{\text{cooling,max}} = \frac{T_{\text{out}}}{T_{\text{out}} - T_{\text{in}}} - 1$$

$$= \frac{T_{\text{in}}}{T_{\text{out}} - T_{\text{in}}} \quad \text{(reversible heat pump)}. \tag{21.29}$$

For a refrigerator, the coefficient of performance of cooling is usually greater than 1 because the temperature of the interior T_{in} is around 255 K, and the temperature difference $T_{\text{out}} - T_{\text{in}}$ between the outside air and the interior is around 40 K.

Figure 21.42 Energy input-output diagram for a heat pump.



EXAMPLE 21.7 Cost of cooling off

The thermal insulation in a refrigerator is never perfect. Typically energy is transferred thermally from the room to the refrigerator interior at a rate of 60 W with the refrigerator door closed. Suppose the interior needs to be held at 4.0 °C and the temperature of the room is 20 °C. At what minimum rate must mechanical energy be delivered to the refrigerator to maintain the interior at 4.0 °C?

❶ **GETTING STARTED** Because I am interested in the minimum rate at which mechanical energy must be delivered, I can treat the refrigerator as a reversible heat pump. The room is the high-temperature reservoir, so $T_{out} = 20\ °C = 293\ K$. The refrigerator interior is the low-temperature reservoir, so $T_{in} = 4.0\ °C = 277\ K$. To maintain the temperature of the interior, the energy Q_{in} that leaks through the thermal insulation and moves from the room into the refrigerator must be transferred out. I am given that the rate at which this energy is transferred thermally into the refrigerator is $dQ_{in}/dt = 60\ W$.

❷ **DEVISE PLAN** To calculate the rate at which mechanical energy must be delivered to balance a given quantity Q_{in} moving from the room to the refrigerator interior, I should first calculate how much work needs to be done to transfer that quantity of energy out of the refrigerator. To calculate the work required to transfer this quantity of energy Q_{in} I can use Eq. 21.27: $W = Q_{in}/COP_{cooling}$. This expression tells me that the amount of work needed is minimum when $COP_{cooling}$ has its maximum value, which I can determine from Eq. 21.29. To obtain the rate at which mechanical energy must be delivered, I take the time derivative of W.

❸ **EXECUTE PLAN** The maximum coefficient of performance of cooling is (Eq. 21.29)

$$COP_{cooling,max} = \frac{277\ K}{293\ K - 277\ K} = 17,$$

so the minimum work required is $W_{min} = Q_{in}/COP_{cooling,max}$. Taking the derivative of this expression, I obtain

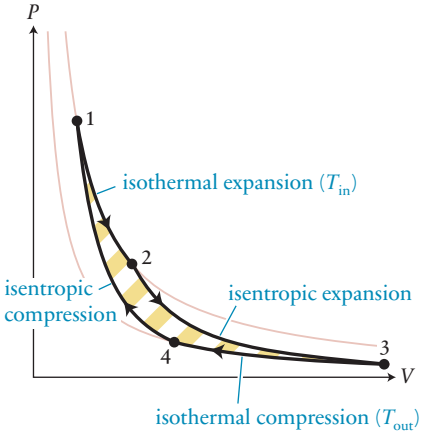
$$\left(\frac{dW}{dt}\right)_{min} = \frac{dQ_{in}/dt}{COP_{cooling,max}} = \frac{60\ W}{17} = 3.5\ W. \checkmark$$

❹ **EVALUATE RESULT** My calculated value for the rate at which mechanical energy must be delivered to cancel the thermal energy that leaks from the room to the refrigerator interior is 1/17 the rate at which that leaking occurs. This makes sense because the coefficient of performance of cooling is 17. This rate is small for what I expect for typical electrical appliances (several hundred watts), but I assumed the refrigerator was reversible. In reality, it is not reversible, and additional energy is dissipated, which means my calculated value is too low. Also, refrigerator doors are opened many times a day, thus increasing the rate of energy transferred thermally to the interior.

21.20 (*a*) If the heat pump from the refrigerator in Example 21.7 is used during the winter to heat a room to 20 °C when it's 4 °C outside, what is $COP_{heating}$? (*b*) At what rate must work be done on the heat pump if you want to heat the room at a rate of 500 W?

21.7 Carnot cycle

Figure 21.43 *PV* diagram for a Carnot cycle.



In Section 21.4, we identified the four processes that make up the Carnot cycle: two isothermal processes and two isentropic processes (**Figure 21.43**). Part of the energy taken in from a high-temperature reservoir is upgraded to mechanical energy, while the rest is degraded and discarded into a low-temperature reservoir (Figure 21.32*b*). The thermal transfer of energy occurs only during the isothermal processes at the two temperatures T_{in} and T_{out}.

As we study the Carnot cycle further, let us use an ideal gas as the working substance so that we can use our results for the entropy and thermal transfer of energy for each process from Tables 20.6 and 20.7, which are gathered together in **Table 21.1**.

To calculate the efficiency of a heat engine operating on this cycle from Eq. 21.21, we need to determine W and Q_{in}. We can obtain the work done on the working substance (our system) during the cycle from Table 21.1:

$$W = -Nk_B T_{in} \ln\left(\frac{V_2}{V_1}\right) - Nk_B T_{out} \ln\left(\frac{V_4}{V_3}\right) \quad \text{(Carnot cycle)}. \qquad (21.30)$$

We do not need to write the two isentropic work terms because they cancel out.

QUANTITATIVE TOOLS

Table 21.1 Entropy and energy changes for the working substance in a Carnot cycle (see Tables 20.6 and 20.7)

Process	Definition	ΔS_{dev}	ΔE_{th}	W	Q
$1 \rightarrow 2$ (*isothermal*)	$\Delta T = 0$	$N \ln\left(\frac{V_2}{V_1}\right)$	0	$-Nk_{\text{B}}T_{\text{in}} \ln\left(\frac{V_2}{V_1}\right)$	$Nk_{\text{B}}T_{\text{in}} \ln\left(\frac{V_2}{V_1}\right)$
$2 \rightarrow 3$ (*isentropic*)	$Q = 0$	0	$NC_V(T_{\text{out}} - T_{\text{in}})$	$NC_V(T_{\text{out}} - T_{\text{in}})$	0
$3 \rightarrow 4$ (*isothermal*)	$\Delta T = 0$	$N \ln\left(\frac{V_4}{V_3}\right)$	0	$-Nk_{\text{B}}T_{\text{out}} \ln\left(\frac{V_4}{V_3}\right)$	$Nk_{\text{B}}T_{\text{out}} \ln\left(\frac{V_4}{V_3}\right)$
$4 \rightarrow 1$ (*isentropic*)	$Q = 0$	0	$NC_V(T_{\text{in}} - T_{\text{out}})$	$NC_V(T_{\text{in}} - T_{\text{out}})$	0
cycle	reversible	0	0	$-Nk_{\text{B}}(T_{\text{in}} - T_{\text{out}}) \ln\left(\frac{V_2}{V_1}\right)$	$-W$

For a complete cycle the entropy change must be zero, and so summing the entropy changes in Table 21.1, we get

$$N \ln\left(\frac{V_2}{V_1}\right) + N \ln\left(\frac{V_4}{V_3}\right) = 0 \tag{21.31}$$

or

$$\ln\left(\frac{V_2}{V_1}\right) = -\ln\left(\frac{V_4}{V_3}\right), \tag{21.32}$$

which means that

$$\frac{V_2}{V_1} = \frac{V_3}{V_4} \quad \text{(Carnot cycle)}. \tag{21.33}$$

Substituting Eq. 21.32 into Eq. 21.30, we obtain

$$W = -Nk_{\text{B}}(T_{\text{in}} - T_{\text{out}}) \ln\left(\frac{V_2}{V_1}\right) \quad \text{(Carnot cycle)}. \tag{21.34}$$

Because the volume ratio V_2/V_1 is greater than 1, $\ln(V_2/V_1)$ is positive. Because $T_{\text{in}} - T_{\text{out}}$ is also positive, the work done on the system is negative, which is consistent with our observation in Section 21.4: The work done on the system for a clockwise cycle is negative. Hence for a clockwise Carnot cycle, the work done on the environment is positive.

The high-temperature reservoir transfers energy thermally to the working substance during the first isothermal process (the one at T_{in}). The quantity Q_{in} is thus equal to the energy transferred thermally during the first isothermal process (see Table 21.1):

$$Q_{\text{in}} = Q_{1\rightarrow 2} = Nk_{\text{B}}T_{\text{in}} \ln\left(\frac{V_2}{V_1}\right). \tag{21.35}$$

Remember that not all of this energy can be converted to mechanical energy because some of it must be transferred to the low-temperature reservoir during the second isothermal process to satisfy the entropy law (Figure 21.32*b*).

Substituting Eqs. 21.34 and 21.35 into Eq. 21.21 thus yields

$$\eta = \frac{-W}{Q_{\text{in}}} = \frac{Nk_{\text{B}}(T_{\text{in}} - T_{\text{out}}) \ln\left(\frac{V_2}{V_1}\right)}{Nk_{\text{B}}T_{\text{in}} \ln\left(\frac{V_2}{V_1}\right)} = \frac{T_{\text{in}} - T_{\text{out}}}{T_{\text{in}}} \quad \text{(Carnot cycle)}, \tag{21.36}$$

which is exactly the maximum efficiency we obtained in Eq. 21.23 for any heat engine, regardless of the working substance.

EXAMPLE 21.8 Carnot engine

In a reversible heat engine, 6.02×10^{23} molecules of nitrogen gas are used as the working substance. The gas undergoes a Carnot cycle as illustrated in Figure 21.43. The volume of the gas is 6.00×10^{-3} m^3 in state 1 and 14.0×10^{-3} m^3 in state 2. The low-temperature reservoir is at 320 K. How much energy is transferred thermally to the low-temperature reservoir during one cycle?

❶ **GETTING STARTED** The temperature of the low-temperature reservoir is $T_{out} = 320$ K. I am also given that $V_1 = 6.00 \times 10^{-3}$ m^3 and $V_2 = 14.0 \times 10^{-3}$ m^3. I need to determine Q_{out}, the amount of energy transferred thermally to the low-temperature reservoir in each cycle.

❷ **DEVISE PLAN** From Eq. 21.2 I know that $Q_{out} = W + Q_{in}$. Therefore this problem boils down to determining W and Q_{in}, which are given by Eqs. 21.34 and 21.35, respectively. These equations both contain T_{in}, which is not given, but the terms containing T_{in} cancel and so T_{in} drops out.

❸ **EXECUTE PLAN** Substituting Eqs. 21.34 and 21.35 into Eq. 21.2 after solving the latter for Q_{out}, I get

$$Q_{out} = W + Q_{in} = -Nk_B(T_{in} - T_{out}) \ln\left(\frac{V_2}{V_1}\right) + Nk_B T_{in} \ln\left(\frac{V_2}{V_1}\right) = Nk_B T_{out} \ln\left(\frac{V_2}{V_1}\right). \quad (1)$$

Substituting the values given yields

$$Q_{out} = (6.02 \times 10^{23})(1.38 \times 10^{-23}\ \text{J/K})(320\ \text{K}) \times \ln\left(\frac{14.0 \times 10^{-3}\ \text{m}^3}{6.00 \times 10^{-3}\ \text{m}^3}\right) = 2.25 \times 10^3\ \text{J}. \checkmark$$

❹ **EVALUATE RESULT** The quantity I obtain for Q_{out} is positive, as it should be. I also note that the expression I obtain for Q_{out} in Eq. 1 can be written as $-Nk_B T_{out} \ln(V_4/V_3)$ because $\ln(V_2/V_1) = -\ln(V_4/V_3)$ (Eq. 21.32). Therefore $Q_{out} = -Q_{3\to4}$ (Table 21.1). This makes sense because energy is transferred thermally to the low-temperature reservoir only during the second isothermal process ($3 \to 4$). The minus sign in $-Q_{3\to4}$ accounts for the fact that $Q < 0$ (energy transferred out of the system) and, by definition, Q_{out} is always positive.



21.21 What is the change in entropy during each of the four legs of the cycle in Example 21.8?

21.8 Brayton cycle

As we saw in Section 21.4, the Brayton cycle (**Figure 21.44**) produces more mechanical energy than the Carnot cycle. More importantly, because the Brayton cycle does not involve any isothermal processes, the rate at which work can be done is greater (yielding more power).

Figure 21.44 *PV* diagram for a Brayton cycle.

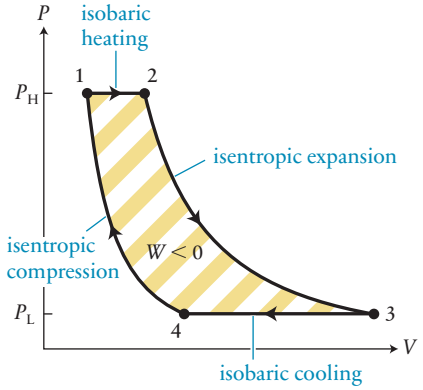


Table 21.2 Entropy and energy for the working substance in a Brayton cycle

Process	Definition	ΔS_{dev}	ΔE_{th}	W	Q
$1 \to 2$ (*isobaric*)	$\Delta P = 0$	$\frac{NC_P}{k_B}\ln\left(\frac{T_2}{T_1}\right)$	$NC_V(T_2 - T_1)$	$-Nk_B(T_2 - T_1)$	$NC_P(T_2 - T_1)$
$2 \to 3$ (*isentropic*)	$Q = 0$	0	$NC_V(T_3 - T_2)$	$NC_V(T_3 - T_2)$	0
$3 \to 4$ (*isobaric*)	$\Delta P = 0$	$\frac{NC_P}{k_B}\ln\left(\frac{T_4}{T_3}\right)$	$NC_V(T_4 - T_3)$	$-Nk_B(T_4 - T_3)$	$NC_P(T_4 - T_3)$
$4 \to 1$ (*isentropic*)	$Q = 0$	0	$NC_V(T_1 - T_4)$	$NC_V(T_1 - T_4)$	0
cycle	reversible	0	0	$-Q$	$NC_P(T_2 - T_1 + T_4 - T_3)$

To determine the efficiency of the Brayton cycle, we begin by collecting the appropriate values for the entropy and energy changes from Tables 20.6 and 20.7. The result is shown in **Table 21.2**.

Because the energy change over a cycle is zero, we have $W = -Q$. To obtain the work done during a cycle, therefore, we can sum the two terms in the rightmost column of Table 21.2 (which is simpler than taking the sum of the four W terms):

$$W = -Q = -NC_P(T_2 - T_1 + T_4 - T_3) \quad \text{(Brayton cycle)}. \quad (21.37)$$

To calculate the efficiency (Eq. 21.21; $\eta = -W/Q_{\text{in}}$), we must obtain expressions for $-W$ and Q_{in}, both of which must be positive. Because $T_2 > T_1$, Q is positive during the first isobaric process ($1 \to 2$) at the high pressure P_H. Likewise, because $T_4 < T_3$, Q is negative during the second isobaric process ($3 \to 4$) at the low pressure P_L. The energy transferred thermally from the environment to the working substance is thus equal to the energy transferred thermally during the isobaric process $1 \to 2$:

$$Q_{\text{in}} = NC_P(T_2 - T_1). \quad (21.38)$$

Substituting Eqs. 21.37 and 21.38 into Eq. 21.21 and dividing out the factor NC_P, we obtain

$$\eta \equiv \frac{-W}{Q_{\text{in}}} = \frac{T_2 - T_1 + T_4 - T_3}{T_2 - T_1} = 1 - \frac{T_3 - T_4}{T_2 - T_1}. \quad (21.39)$$

We can simplify this expression for the efficiency of the Brayton cycle by using our knowledge about entropy. The entropy change for the cycle is zero (Table 21.2). This ΔS_{dev} for the cycle is the sum of the two nonzero ΔS_{dev} terms in Table 21.2, so setting the sum of these two terms equal to zero gives

$$\frac{NC_P}{k_B}\ln\left(\frac{T_2}{T_1}\right) + \frac{NC_P}{k_B}\ln\left(\frac{T_4}{T_3}\right) = 0 \quad (21.40)$$

$$\ln\left(\frac{T_2}{T_1}\right) = -\ln\left(\frac{T_4}{T_3}\right) = \ln\left(\frac{T_3}{T_4}\right), \quad (21.41)$$

which means that

$$\frac{T_2}{T_1} = \frac{T_3}{T_4} \quad \text{(Brayton cycle)}. \quad (21.42)$$

We now rewrite Eq. 21.39 as

$$\eta = 1 - \frac{T_4}{T_1}\left[\frac{(T_3/T_4) - 1}{(T_2/T_1) - 1}\right]. \tag{21.43}$$

By Eq. 21.42, the term inside the square brackets is equal to 1, so the efficiency is

$$\eta = 1 - \frac{T_4}{T_1} \quad \text{(Brayton cycle)}. \tag{21.44}$$

The temperature T_4 is the lowest temperature in the cycle. The temperature T_1 is the temperature of the working substance at the completion of the isentropic compression, when the working substance exits the compressor (shown in Figure 21.30*b*).

During the isentropic compression, we have, from Eq. 20.47,

$$(P_H)^{(1/\gamma)-1}T_1 = (P_L)^{(1/\gamma)-1}T_4, \tag{21.45}$$

where γ is the heat capacity ratio C_P/C_V given by Eq. 20.26. The temperature ratio in Eq. 21.44 then becomes

$$\frac{T_4}{T_1} = \left(\frac{P_H}{P_L}\right)^{(1/\gamma)-1}. \tag{21.46}$$

The efficiency in Eq. 21.44 can thus be expressed in terms of the pressure increase accomplished by the compressor in the process $4 \rightarrow 1$:

$$\eta = 1 - \left(\frac{P_H}{P_L}\right)^{(1/\gamma)-1} \quad \text{(Brayton cycle)}. \tag{21.47}$$

The ratio P_H/P_L is called the *pressure ratio*. For aircraft engines, this ratio typically is between 4 and 30. The higher the pressure ratio of an engine running on the Brayton cycle, the more efficient the engine.

EXAMPLE 21.9 Performance and pressure ratio

Consider a jet engine operating on the Brayton cycle. The working substance is air, which consists primarily of diatomic molecules. The air is drawn in at atmospheric pressure, 1.01×10^5 Pa, at 288 K. In the combustion chamber, the temperature of the air is raised to 1520 K, and the engine has a pressure ratio of 30. What are the temperatures at the end of the two isentropic processes?

❶ **GETTING STARTED** Air enters the engine at point 4 in the cycle in Figure 21.44. Therefore I know that $T_4 = 288$ K. The air is compressed between state 4 and state 1 and the combustion takes place between state 1 and state 2, so I also know that $T_2 = 1520$ K. The isentropic processes end at states 1 and 3, so I need to determine T_1 and T_3.

❷ **DEVISE PLAN** To determine T_1 I can use Eq. 21.46, which relates T_1 to T_4 and the pressure ratio, both of which are given. This equation also contains the heat capacity ratio γ, which I can determine from Eq. 20.26, $\gamma = C_P/C_V = 1 + 2/d$, because I know that $d = 5$ for diatomic molecules. Once I have T_1, I can use Eq. 21.42 to determine T_3.

❸ **EXECUTE PLAN** From Eq. 20.26, I get

$$\gamma = 1 + \frac{2}{5} = \frac{7}{5},$$

so the exponent in Eq. 21.46 is $\frac{1}{\gamma} - 1 = \frac{5}{7} - 1 = -\frac{2}{7}$. I can then solve Eq. 21.46 for T_1:

$$T_1 = \frac{T_4}{(P_H/P_L)^{(1/\gamma)-1}} = \frac{288\text{ K}}{30^{-2/7}} = 760\text{ K}. \checkmark \tag{1}$$

Now that I have T_1 and T_2, I can get T_3 from Eq. 21.42:

$$T_3 = \frac{T_2 T_4}{T_1} = \frac{(1520\text{ K})(288\text{ K})}{760\text{ K}} = 575\text{ K}. \checkmark$$

❹ **EVALUATE RESULT** I know from Section 20.4 that on a given PV diagram, isentropes are steeper than isotherms, so Figure 21.44 tells me that T_1 and T_3 should lie between $T_4 = 288$ K and $T_2 = 1520$ K. The values I obtain indeed satisfy this condition.

21.22 (*a*) What is the efficiency of the jet engine in Example 21.9? (*b*) If the jet engine were to operate on a Carnot cycle between the same highest and lowest temperatures, what would its efficiency be?

QUANTITATIVE TOOLS

Chapter Summary

Converting energy (Sections 21.1, 21.5)

Concepts A system's **energy input** is the sum of the mechanical input of energy W_{in} and the thermal input of energy Q_{in}. The **energy output** is the sum of the mechanical output of energy W_{out} and the thermal output of energy Q_{out}.

The work done on a system is

$$W = W_{in} - W_{out},$$

and the energy transferred thermally to the system is

$$Q = Q_{in} - Q_{out}.$$

A **steady device** converts energy in such a way that the energy changes and entropy changes of the device are zero either continuously or over some repeating time interval (a *cycle*).

Quantitative Tools For a steady device over one cycle:

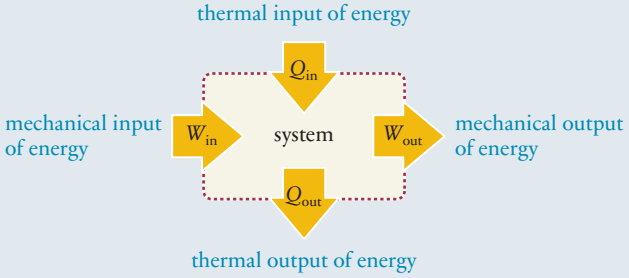
$$\Delta E = 0 \tag{21.1}$$

$$W = Q_{out} - Q_{in} \tag{21.2}$$

$$\Delta S_{dev} = 0 \tag{21.3}$$

$$\Delta S_{env} \geq 0. \tag{21.5}$$

Figure SUM21.1



Entropy and energy quality (Sections 21.2, 21.5)

Concepts An **entropy diagram for a steady device** consists of an entropy gradient axis, one or more rectangles representing energy transfers from input to output, and downward- or upward-pointing arrows representing thermal or mechanical inputs or outputs of energy. Each rectangle has an arrowhead to indicate the direction of energy transfer. Thermal energy transferred into the device causes the entropy of the environment to decrease; thermal energy transferred out of the device causes the entropy of the environment to increase. (See the Procedure box "Drawing entropy diagrams for a steady device.")

A *degradation* in an environment's **energy quality** is associated with an increase in the environment's entropy; an *upgrading* in an environment's energy quality is associated with a decrease in the environment's entropy. The higher the quality of energy transferred, the smaller the associated entropy change. The amount of useful work obtainable from a quantity of energy decreases as the quality of the energy decreases.

Quantitative Tools The **entropy gradient** dS/dE of a system is the rate at which the entropy of the system changes with a change in thermal energy:

$$\frac{dS}{dE} \equiv \frac{1}{k_B T}.$$

See the Procedure box "Computing entropy changes from entropy diagrams" for how to use an entropy diagram to compute an entropy change.

Heat engines and heat pumps (Sections 21.3, 21.6)

Concepts A **heat engine** does work on its environment by thermally taking in energy Q_{in} at a high temperature and thermally expelling energy Q_{out} at a lower temperature.

A **heat pump** uses mechanical energy to take in energy thermally at a low temperature and transfer out energy thermally at a higher temperature.

Quantitative Tools For a reversible heat engine in which Q_{in} enters at temperature T_{in} and Q_{out} leaves at T_{out},

$$\frac{Q_{out}}{Q_{in}} = \frac{T_{out}}{T_{in}}. \tag{21.20}$$

The **efficiency** η of a heat engine is

$$\eta \equiv -\frac{W}{Q_{in}} = 1 - \frac{Q_{out}}{Q_{in}}, \tag{21.21, 21.22}$$

where W is the work done on the steady device during one cycle. The maximum efficiency is achieved by a reversible heat engine and is

$$\eta_{max} = 1 - \frac{T_{out}}{T_{in}}. \tag{21.23}$$

The **coefficient of performance of heating** for a heat pump is

$$\text{COP}_{heating} \equiv \frac{Q_{out}}{W}. \tag{21.24}$$

The **coefficient of performance of cooling** for a heat pump is

$$\text{COP}_{cooling} \equiv \frac{Q_{in}}{W} = \text{COP}_{heating} - 1. \tag{21.27, 21.28}$$

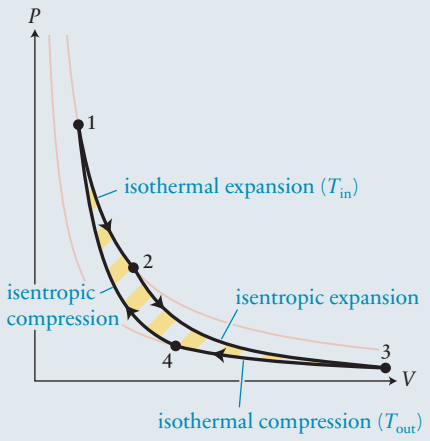
Thermodynamic cycles (Sections 21.4, 21.7, 21.8)

Concepts The magnitude of the work done on a system during one thermodynamic cycle is equal to the area enclosed by the path that represents the cycle in a PV diagram. The work done on the system is negative for a clockwise cycle and positive for a counterclockwise cycle.

The **Carnot cycle** consists of two isothermal processes and two isentropic processes. It operates at the maximum possible efficiency between two temperature reservoirs.

Quantitative Tools

Figure SUM21.2

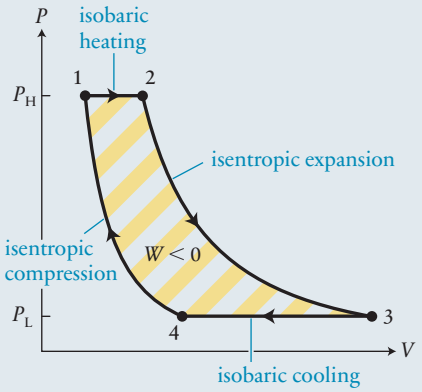


The **Brayton cycle** consists of two isobaric processes and two isentropic processes.

The efficiency of a Carnot engine is

$$\eta = \frac{T_{in} - T_{out}}{T_{in}}. \tag{21.36}$$

Figure SUM21.3



The efficiency of a Brayton cycle that operates between pressures P_H and P_L is

$$\eta = 1 - \frac{T_4}{T_1} = 1 - \left(\frac{P_H}{P_L}\right)^{\frac{1}{\gamma}-1}, \tag{21.44, 21.47}$$

where T_1 is the temperature of the working substance when it exits the compressor, T_4 is the lowest temperature in the cycle, and γ is C_P/C_V. The quantity P_H/P_L is the *pressure ratio* of the cycle.

Summary of Tables 21.1 and 21.2 (from Tables 20.6 and 20.7)

Process	W	Q	ΔE_{th}	ΔS
isochoric	0	$NC_V\Delta T$	Q	$\frac{NC_V}{k_B}\ln\left(\frac{T_f}{T_i}\right)$
isothermal	$-Nk_BT\ln\left(\frac{V_f}{V_i}\right)$	$-W$	0	$N\ln\left(\frac{V_f}{V_i}\right) = \frac{Q}{k_BT}$
isobaric	$-Nk_B\Delta T$	$NC_P\Delta T$	$NC_V\Delta T$	$\frac{NC_P}{k_B}\ln\left(\frac{T_f}{T_i}\right)$
isentropic	ΔE_{th}	0	$NC_V\Delta T$	0

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

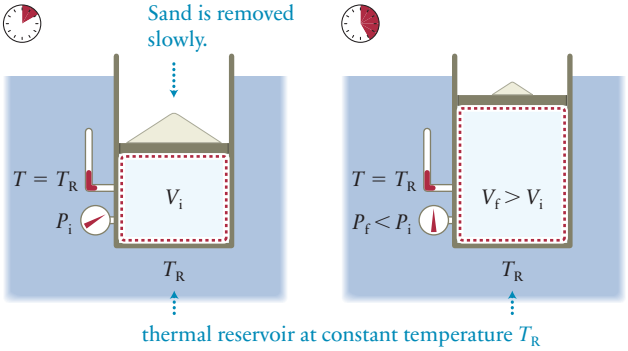
Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

Unless stated otherwise, take the working substance in all heat engines, heat pumps, and thermodynamic cycles to be an ideal gas.

21.1 Converting energy

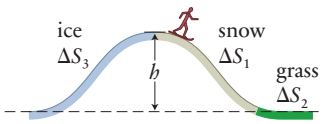
1. Figure P21.1 shows a gas confined to a cylinder fitted with a movable piston. The cylinder is immersed in a tank of water that acts as a thermal reservoir, so that the temperature of the gas never changes. (*a*) With the system defined as the gas, draw an energy input-output diagram for the system as sand grains are slowly removed from the top of the piston and the piston rises. (*b*) Does the system's energy change? Does its entropy change? (*c*) Is this a steady device? •

Figure P21.1



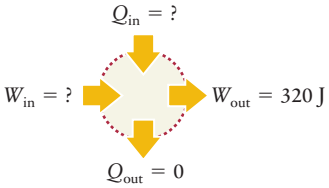
2. Draw an energy input-output diagram for a steady device that partially converts thermal energy to work, drawing your input and output arrows such that their relative sizes indicate the relative amounts of the quantities being represented. Label changes in the energy and entropy of the device. •
3. The snowboarder in Figure P21.3 starts at the top of a snow-covered hill and boards down the hill, which produces a change in entropy ΔS_1. At the bottom, she slides over a patch of grass, which quickly slows her down and produces a change in entropy ΔS_2. She then climbs back up and slides down the other side of the hill, which is covered in ice rather than snow, producing a change in entropy ΔS_3. Rank the entropy changes by magnitude, smallest first. •

Figure P21.3



4. For the steady device shown in Figure P21.4, determine the values of W_{in} and Q_{in}. ••

Figure P21.4



5. A steady heat engine, in contact with its environment, absorbs energy of magnitude Q_H as heat from a high-temperature reservoir, while 120 J of work is done on it. Without specifying anything about the way the entropy of the engine, source, or environment responds to change in energy, find the value of Q_H when the maximum amount of work the engine can do is arbitrarily close to 375 J. •
6. A 1500-kg car constantly moving at 90 kmph accidentally crashes into a concrete wall. The impact produces 750 J of sound energy. The car gains 3.8 kJ of thermal energy from the impact, and 15 kJ of thermal energy escapes after the impact. How much energy does the concrete wall absorb? ••
7. A steady device operates on the 600-W output of a motor. The device runs at 20 Hz and can convert 68% of its input to usable power to pump water. (*a*) How much thermal energy does the device output per cycle? (*b*) How much thermal energy does the device output in 8.0 h? (*c*) If 50% of the output thermal energy of part *b* were captured and recycled, how many kilograms of water could it raise from ground level to a height of 10 m? ••

21.2 Quality of energy

8. Three reservoirs of heated water are represented on the entropy diagram in Figure P21.8. All three reservoirs contain equal amounts of energy, so that $E_1 = E_2 = E_3$. (*a*) For which reservoir would you be least concerned about burning yourself if you dunked your hand into the water? (*b*) Which reservoir is most useful for converting thermal energy to mechanical energy? •

Figure P21.8

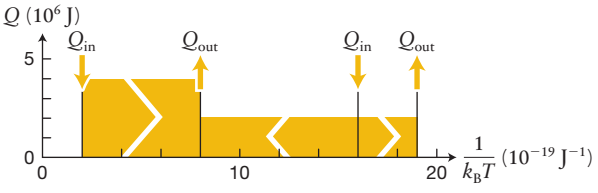


9. The entropy of material A varies as $S_1 = -4\gamma/E$, where γ is a constant, and the entropy of material B varies as $S_2 = \gamma E^2/4$. Which material is better suited to converting higher-quality thermal energy to useful, higher-quality mechanical energy when $E = 2$ units of energy? ••
10. A reversible steady device operates between two thermal reservoirs, one at 390 K and one at 250 K. If no work is done on the device, what is the maximum value for the percentage of thermal energy input Q_{in} converted to mechanical energy W_{out}? (Hint: Consider the entropy diagram of a reversible process and the change in entropy.) ••
11. Find the temperature in Kelvin and degree Celsius of a thermal reservoir with an entropy gradient value of $1.985 \times 10^{20}\ \text{J}^{-1}$. •
12. The entropy of a thermal reservoir of unknown temperature is defined by the function $S = aE^3/3b^2$. (*a*) Derive an expression for the entropy gradient dS/dE. (*b*) Calculate the entropy gradient of the reservoir if the ratio E/b is $3.00\ \text{J}^{-1}$ and $a = 1.00$ J. (*c*) Find the temperature of the reservoir. ••

21.3 Heat engines and heat pumps

13. Draw an entropy diagram for a reversible steady device that takes seawater at 29 °C and upgrades as much energy as possible to a reservoir filled with water at 100 °C and on the verge of boiling. The rest of the energy is deposited at 14 °C, the temperature of deeper seawater. •
14. A composite of several processes that take place in a steady device is represented by the entropy diagram in Figure P21.14. What is ΔS for the composite? •

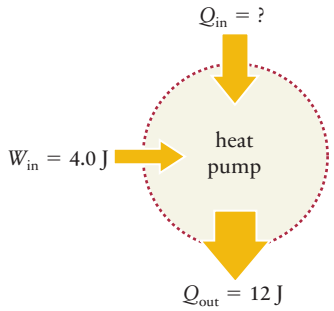
Figure P21.14



15. A reversible steady device puts out $W_{out} = 750$ J of mechanical energy. The only input energy is the quantity Q_{in} transferred thermally from a thermal reservoir at 355 K. If the temperature of the environment (which serves as a low-temperature reservoir) is 300 K, what quantity of energy Q_{out} is transferred thermally out of the device? ••

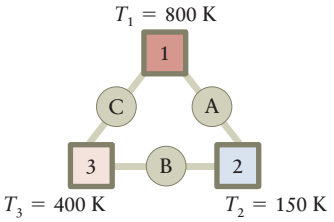
16. The energy input-output diagram for a heat pump is shown in Figure P21.16. What are (*a*) the value of Q_{in}, (*b*) the coefficient of performance of heating, and (*c*) the coefficient of performance of cooling? ••

Figure P21.16



17. Three identical steady devices, A, B, and C, operate between three thermal reservoirs, 1, 2, and 3, as shown in Figure P21.17. Each device takes thermal energy from the warmer of its two reservoirs, and each upgrades to mechanical energy as much of that thermal energy as possible without lowering the entropy of the three-device, three-reservoir system. (*a*) Which reservoir has the highest-quality energy? (*b*) Is enough information provided to determine which device can upgrade the highest percentage of its input thermal energy to mechanical energy? If so, which is the most efficient of the three devices? If not, what information is lacking? •••

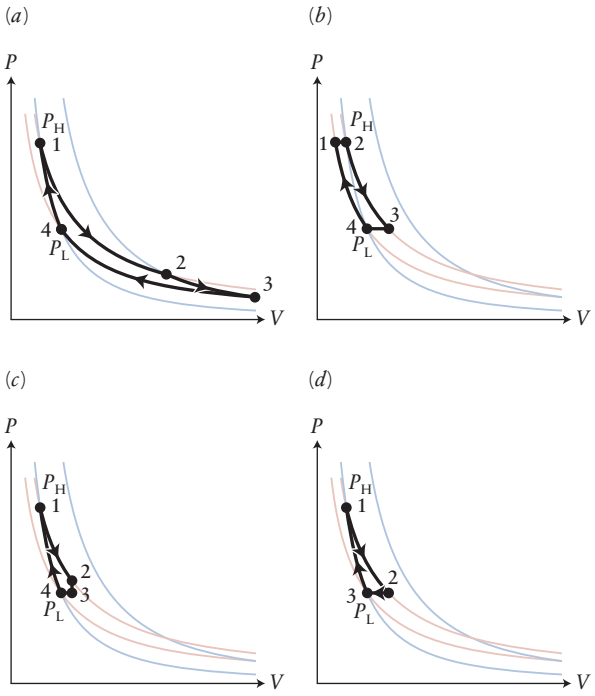
Figure P21.17



21.4 Thermodynamic cycles

18. Figure P21.18 shows *PV* diagrams for four thermodynamic cycles. In each case, the cycle consists of three or four processes carried out on identical working substances, and each process is isothermal, isobaric, isochoric, or isentropic. The pressures P_H and P_L have the same values in all four cycles. For which single cycle is the most work done on the gas? •

Figure P21.18



19. Which of the four thermodynamic cycles illustrated in Figure P21.18 has the greatest efficiency? ••

20. Two thermodynamic processes, each involving the same type and amount of monatomic gas as a working substance and each working as a heat engine, begin and end their cycles from a common state: $P_1 = 1.00 \times 10^5$ Pa, $V_1 = 2.50 \times 10^{-2}$ m^3, and $T_1 = 300$ K. The first is a Carnot cycle, the second is a Brayton cycle, and each does 200 J of work on the environment during each cycle. (*a*) Draw an entropy diagram for each process. (*b*) Which process causes the greater change in entropy of the environment? ••

21.5 Entropy constraints on energy transfers

21. The working substance in a steady device is a gas that undergoes expansion followed by compression. The work done by the gas during expansion is 6.4 J, and the work done on it during compression is 8.2 J. What is the work done on the gas during the cycle? •

22. What is the entropy change in the environment when 5.0 MJ of energy is transferred thermally from a reservoir at 1000 K to one at 500 K? •

23. In one cycle, a steady device transfers 1.55×10^6 J of thermal energy from a reservoir at 450 K to a reservoir at 300 K. Calculate the entropy change for the device and for the environment. •

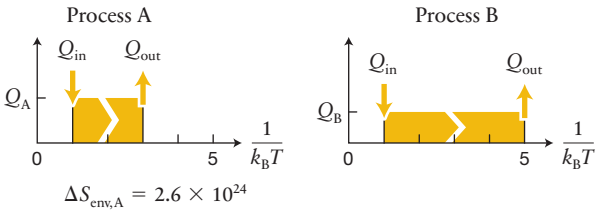
24. Draw an entropy diagram for a steady device that in each cycle converts 135 J of work to thermal energy that is released into a reservoir at 340 K. Calculate numerical values for the length and height of the entropy bar and for the change in entropy that results from this transfer. ••

25. An ideal gas is held at 70 K in a cylinder fitted with a movable piston. When a 5.0-kg cube is placed on the piston, the piston slides down a distance of 200 mm. What

is the entropy change in the environment that results from the change in the gas volume? ••

26. Processes A and B in Figure P21.26 each transfer energy thermally. If the energies transferred thermally by the two processes are related by $Q_A = \sqrt{2} \times Q_B$, and process A causes $\Delta S_{env} = 2.6 \times 10^{24}$, what is ΔS_{env} for process B? ••

Figure P21.26



21.6 Heat engine performance

27. A reversible heat engine transfers energy from a 426-K reservoir to second reservoir. What must be the temperature of the second reservoir if the engine has a maximum efficiency of 0.350? •

28. A heat engine takes in 6.45×10^3 J of thermal energy from a reservoir at 500 K and returns some of this energy to a reservoir at $T_L < 500$ K. If this engine operates at an efficiency of 0.450, what is the maximum value possible for T_L? •

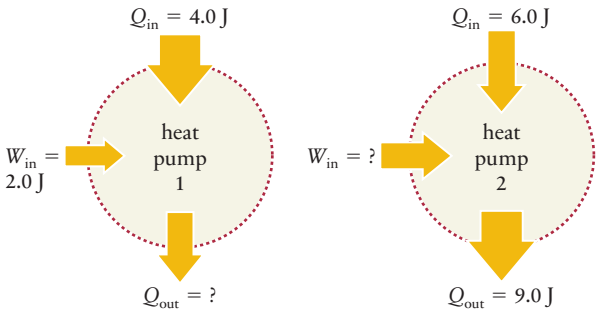
29. What is the coefficient of performance of heating of a heat engine that takes in 6.45 kJ and expels 7.30 kJ of thermal energy? •

30. A heat pump with a coefficient of performance of heating of 14 supplies thermal energy to a house to maintain a comfortable temperature for the occupants. Calculate the daily thermal energy supplied to the house if 6.8 kW of power is needed to drive the heat pump. ••

31. A heat engine does 85 J of work per cycle while expelling 110 J of waste thermal energy. How much energy is transferred thermally to this engine each cycle, and what is the efficiency of the engine? ••

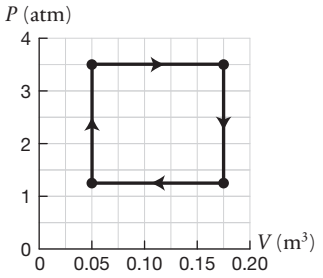
32. Energy input-output diagrams for two heat pumps are shown in Figure P21.32. What are (*a*) Q_{out} for pump 1 and (*b*) W_{in} for pump 2? ••

Figure P21.32



33. The heat engine represented by the *PV* diagram in Figure P21.33 exhausts 43.5 kJ of thermal energy per cycle. What is the engine's efficiency? ••

Figure P21.33



34. During each cycle, a heat engine ejects 75 J of thermal energy for every 115 J of input thermal energy. This engine is used to lift a 375-kg load a vertical distance of 27.0 m at a steady rate of 52.5 mm/s. How many cycles of the engine are needed to accomplish this task? ••

35. In winter, you like to keep your house interior at 21.0 °C. Your geothermal heating system, which was advertised as being reversible, draws thermal energy from an underground reservoir at 347 K. In a cold winter, with the average outdoor temperature being 0.0 °C, thermal energy escapes from the house at a rate of 1000 W ± 5%. Distributing the thermal energy supplied by the heating system throughout the house requires 180 W (provided by the heating system). Is your heating system reversible within your 5% margin of error, or was the advertisement false? Ignore any inefficiency in converting electrical energy to mechanical energy. ••

36. An engine has a rated maximum efficiency of 0.22 when burning a certain fuel, but the manufacturer claims that burning a new type of fuel will increase the maximum efficiency. In testing this claim, you determine that, with an identical energy input, the ratio of the work done on the environment using the new fuel to the work done on the environment using the old fuel is 1.14. What is the efficiency using the new fuel? ••

37. For a heat pump that operates on a Carnot cycle, calculate (*a*) the coefficient of performance of cooling when the pump is used to cool a house to 72 °F on a day when the outside air is at 105 °F and (*b*) the coefficient of performance of heating when the pump is used to heat a house to 72 °F on a day when the outside air is at −15 °F. ••

38. Reversible heat engines 1 and 2 are connected to each other in such a way that the output temperature of engine 1 is the input temperature of engine 2, $T_{1out} = T_{2in}$, and the quantity of energy transferred thermally out of engine 1 is the quantity transferred into engine 2, $Q_{1out} = Q_{2in}$. Calculate Q_{2out} if $T_{1in} = 270$ K, $T_{2out} = 422$ K, and $Q_{1in} = 3.70 \times 10^{11}$ J. ••

39. On a day when the outdoor temperature is 35 °C, a walk-in freezer in a butcher shop where there is no air-conditioning must remove 3000 J of thermal energy from its interior each second to maintain an interior temperature of −4 °C. (*a*) What is the maximum coefficient of performance of cooling of the heat pump in the freezer? (*b*) What would this maximum coefficient be if the freezer were installed in the shop's basement, where the air temperature is 19 °C? (*c*) How much energy could the owner save in a day by installing the freezer in the basement? •••

40. The internal temperature of a freezer is 269 K, and the temperature of the cooling coils on the back of the freezer is 325 K. (*a*) What is the maximum coefficient of performance of cooling of the freezer? (*b*) You measure the time interval it takes to freeze a sample of water and estimate that, during each cycle, the freezer absorbs 118 J from its interior. How much thermal energy does the freezer transfer to the environment during each cycle? (*c*) How much work is done on the freezer each cycle? (*d*) You notice that the room in which the freezer is located is warmer than nearby rooms, and you wonder how efficient the freezer is at warming the room compared with the efficiency of a small electric space heater. Estimate the relative heating efficiencies of these two devices. ●●●

21.7 Carnot cycle

41. How much work is done by a Carnot cycle that doubles the volume of 1.00 mol of a monatomic ideal gas during isothermal expansion and triples the initial temperature of 0.0 °C during isentropic compression? ●

42. What is the maximum efficiency of a Carnot heat engine that operates between 485 K and 625 K? ●

43. Two Carnot heat engines have the following values for the temperatures T_{in} and T_{out} of the reservoirs:

Engine 1: $T_{in} = 450$ K and $T_{out} = 350$ K

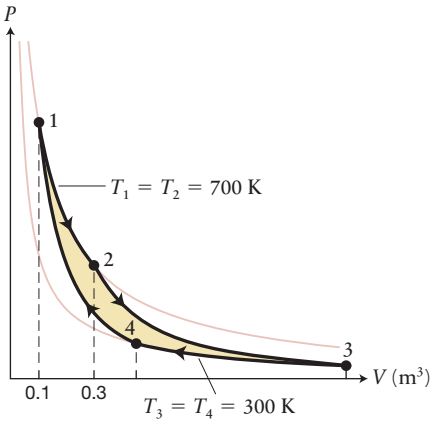
Engine 2: $T_{in} = 550$ K and $T_{out} = 450$ K

Determine which of the two engines is more efficient. ●●

44. A heat engine operates on a Carnot cycle that runs clockwise between a reservoir at 340 K and a reservoir at 280 K. One cycle moves enough energy from the high-temperature reservoir to raise the temperature of 1.0 kg of water by 1.0 K. How much work is done by the engine in one cycle? ●●

45. The Carnot cycle shown in Figure P21.45 uses 1.00 mol of a monatomic ideal gas as its working substance. From the information given in the graph, determine the values of P_1, P_2, P_3, P_4, V_3, and V_4. ●●

Figure P21.45



46. Carnot cycles A and B run on the same amount of the same working substance and draw their thermal energy from the same high-temperature reservoir. The two cycles start with their working substances occupying different volumes. During the isothermal expansion, cycle A doubles its volume and cycle B triples its volume. If cycle A does twice as much work as cycle B, which cycle is more efficient and by how much? ●●

47. Heat engines 1 and 2 operate on Carnot cycles, and the two have the same efficiency. Engine 1 takes in boiling water at 373 K and outputs water that is twice as hot as the output water from engine 2. What is the temperature of the input water for engine 2? ●●

48. A heat pump that operates on a Carnot cycle cools a house in a very hot climate. During the hottest part of the day, the outside temperature is 45 °C, and you want to keep the inside temperature at 20 °C. The heat pump operates on electricity. (*a*) What is the coefficient of performance of cooling during the hottest part of the day? (*b*) At what maximum rate does the pump use electrical energy if it removes thermal energy from the house at a rate of 1.50 kW? ●●

49. A heat pump that operates on a Carnot cycle uses N_2 gas and operates on electricity. (*a*) At what rate is it using electrical energy when it adds thermal energy at a rate of 1.35 kW to a room that is at 22 °C when the outside temperature is −25 °C? (*b*) What is the pump's coefficient of performance of heating? ●●

50. To meet energy regulations, the coefficient of performance of cooling of a refrigerator that operates on a Carnot cycle must be 4.00. At what rate is this refrigerator expelling thermal energy into the kitchen if it consumes electrical energy at a rate of 225 W? ●●

51. A heat engine that operates on a Carnot cycle uses a low-temperature reservoir at 25 °C and runs at an efficiency of 0.350. The working substance is nitrogen gas, and the gas chamber contains 7.50×10^{23} N_2 molecules. What is the change in the thermal energy of the N_2 gas between its lowest and highest temperatures? ●●

52. An electric heat pump used as a heater operates on a Carnot cycle. The low-temperature reservoir consists of a pipe driven 10 m into the ground to a region where the temperature is always 10 °C. The high-temperature reservoir is the interior of a house, kept at 20 °C. (*a*) For every 1.0 J of electrical energy used to run the heat pump, what quantity of thermal energy is delivered to the house? (*b*) If the low-temperature reservoir were the air outside the house at −23 °C rather than underground, what quantity of thermal energy would 1.0 J of electrical energy deliver to the house? Does it seem to be worth the added expense of drilling to tap the underground reservoir? (*c*) What is the coefficient of performance of heating of the heat pump in part *a* and in part *b*? ●●

53. The temperature of Earth's oceans at a depth of 2500 m is about 4 °C. Suppose you want to use this water as the low-temperature thermal reservoir for an electric heat pump to heat the cabin of a boat to 20 °C. If the pump operates on a Carnot cycle and uses electrical energy at a rate of 75 W, at what rate are you changing the entropy of the deep ocean and the entropy of the air in the cabin? ●●

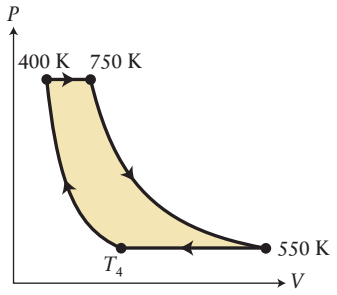
54. An experimental six-cylinder engine that operates on a Carnot cycle has been proposed for a car. Each cylinder is fitted with a movable piston that allows the cylinder to fill with 1.00 mol of a monatomic ideal gas. During isothermal expansion at 700 K, the length of each cylinder increases from 100 mm to 200 mm as the piston moves. The isothermal compression occurs at 400 K. Each isothermal process takes 9.00 s, and each isentropic process takes 2.00 s. If the mass of the car is 750 kg, how long a time interval is required for this car and its 70.0-kg driver to accelerate from rest to 50.0 km/h? Ignore drag and friction. •••

55. A freezer that operates on a Carnot cycle maintains an interior temperature of –18 °C and exhausts thermal energy into a kitchen at 22 °C. You place a tray containing 0.500 kg of water at 0 °C in the freezer and observe that the freezer draws electrical energy at a rate of 115 W. (*a*) How long does it take to transform all the water to ice? (*b*) What is the freezer's coefficient of performance of cooling? (*c*) How much thermal energy is exhausted into the kitchen during the time interval it takes to freeze the water? (*d*) What is the change in entropy of the water as it freezes? (*e*) What is the change in entropy of the system that consists of the water plus the kitchen? For water, the specific transformation energy for melting is $L_{\text{melting}} = 334$ kJ/kg. •••

21.8 Brayton cycle

56. Determine T_4 for the Brayton cycle represented in Figure P21.56. •

Figure P21.56



57. An engine that operates on a Brayton cycle uses oxygen gas (diatomic molecules, O_2) as the working substance. During the isobaric cooling, the temperature decreases by 85 °C. During the isobaric heating, the temperature increases by 100 °C. What is the maximum efficiency of this cycle? •

58. A gas engine that operates on a Brayton cycle has an efficiency of 0.22. On a cold day, the temperature of the air drawn into the engine is 267 K. What is the temperature of the air exhausted from the engine? •

59. An experimental car engine operates on a Brayton cycle and uses a monatomic ideal gas as the working substance. If the pressure in front of the moving car is 103,500 Pa and the pressure behind the moving car is 99,700 Pa, what is the maximum efficiency possible for the engine? ••

60. An engine that operates on a Brayton cycle uses air as its working substance, compressing ambient air initially at a pressure of 1.0 atm to 6.0 atm. (*a*) What is the engine's pressure ratio? (*b*) If the heat capacity ratio $\gamma = C_P/C_V$ is 1.40 for air, what is the engine's efficiency? (*c*) At high altitudes, $\gamma = 1.44$ and the ambient air pressure is 0.40 atm. If the engine compresses this air to 4.8 atm, what is its efficiency now? ••

61. A heat engine that operates on a Brayton cycle has a pressure ratio of 5 and uses as its working substance a gas for which the heat capacity ratio C_P/C_V is 1.333. You wish to replace this engine with one that operates on a Carnot cycle and has the same efficiency as the Brayton engine. If the low-temperature reservoir for the Carnot engine is to be 15 °C, what must the temperature of the high-temperature reservoir be? ••

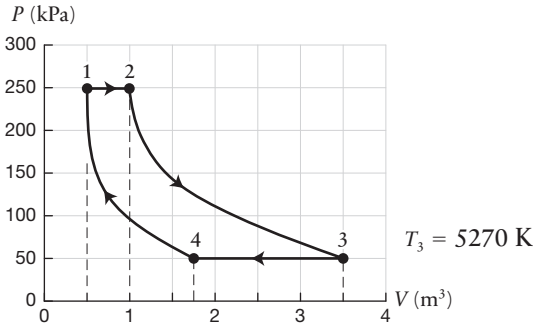
62. You need to replace a heat engine that operates on a Carnot cycle with one that has the same efficiency but operates on a Brayton cycle. The Carnot engine low and high temperatures are −10 °C and 225 °C. If the working substance in your replacement engine has a C_P/C_V ratio of 1.667, what must the engine's pressure ratio be? ••

63. A heat engine that operates on a Brayton cycle uses N_2 gas and has a pressure ratio of 10. It is being used to lift a 535-kg pallet of bricks vertically at a steady rate of 100 mm/s. At what rate must thermal energy be supplied to this engine to accomplish this task? ••

64. You need to design a heat engine that operates on a Brayton cycle and, during each cycle, does 175 J of work while exhausting 65 J of energy. (*a*) If the working substance is helium gas, what should the pressure ratio be? (*b*) If the working substance is carbon dioxide gas, CO_2, is the pressure ratio the same as with helium gas? If so, prove it; if not, determine the new pressure ratio. ••

65. If the Brayton cycle represented in Figure P21.65 operates on 4.0 mol of an ideal gas, how many degrees of freedom do the particles of the gas have? •••

Figure P21.65



66. You are designing a gas engine that operates on a Brayton cycle and uses air ($\gamma = 1.4$) as its working substance. By making the engine larger and larger, you can have the pressure ratio be 16, 17, 18, 19, or 20. At the higher pressure ratios, however, the efficiency is a smaller fraction of the ideal Brayton cycle efficiency. You determine that, over this range of ratios, the factor relating the efficiency of the real engine to that of the ideal cycle, $\eta_{\text{real}} = f\eta_{\text{ideal}}$, can be described by the formula $f = (1.45)(1 - \eta_{\text{ideal}})^{0.8}$. (*a*) Which pressure ratio do

you choose to maximize efficiency? (*b*) This is a simplified problem. Other than increasing mass, what other reasons can you think of for why engine designers don't have the pressure ratio as great as possible? •••

Additional Problems

67. A thermal reservoir is held at 265 K. If 4180 J of energy is transferred thermally to it, what is the change in entropy of the reservoir? •

68. The entropy of a system is increased by 3.8×10^{21} for every 10 J of energy added to it. What is the temperature of this system? •

69. In the past 40 years, the average power a refrigerator requires to provide a given amount of cooling has dropped from 160 W to 40 W. By what factor has the COP of cooling for refrigerators changed during this time interval? •

70. A 1000-kg iceberg at 0 °C falls into the ocean at a location where the water temperature is 2 °C, and all the ice melts by absorbing thermal energy from the water. What is the entropy change for the system comprising the water and the iceberg? For ice, the specific transformation energy for melting is 334 kJ/kg. ••

71. The low-temperature reservoir for a heat engine that operates on a Carnot cycle is at –5 °C. Using only 1.00 MJ of thermal energy, this engine needs to pull a 1200-kg boulder at a steady rate 65.0 m up a ramp inclined at 35.0° above the horizontal. (*a*) What must the temperature of the high-temperature reservoir be? (*b*) How much of the 1.00 MJ of input energy is wasted as the boulder travels this distance? Ignore any friction between the boulder and the ramp surface. ••

72. An electric heat pump that operates on a Carnot cycle is used to cool a house interior to 20 °C on a day when the outside temperature is 38 °C. (*a*) For every 1.0 J of electrical energy used, what quantity of thermal energy is removed from the house? (*b*) What is the pump's coefficient of performance of cooling? ••

73. A Carnot heat engine in a factory operates at an efficiency of 0.33 when it takes in thermal energy from a reservoir of heated water and releases thermal energy into the 28 °C surface water of an outdoor pool. By how much could the efficiency be improved if the thermal energy were released into the pool at a depth where the water temperature is 18 °C? ••

74. A heat engine that operates on a Carnot cycle is used to pump water out of a well that is 45.0 m deep. The high-temperature reservoir for the heat engine is at 215 °C, and the low-temperature reservoir is at −10 °C. The engine must move water from the bottom of the well to the surface at a rate of 6600 L/min, and the speed of the water when it reaches the surface is negligible. At what rate must thermal energy be supplied to the engine? ••

75. An electric heat pump that operates on a Carnot cycle is used to keep the air in a laboratory chamber at a constant 35 °C by extracting energy from outside air that is at 5 °C. Due to conduction and small leaks, the chamber loses thermal energy at a rate of 12.0 W. (*a*) At what rate must the pump extract energy from the outside air? (*b*) At what rate is the pump using electrical energy? ••

76. A heat pump that operates on a Carnot cycle must deliver thermal energy to your house at a rate of 12.5 kW to keep the rooms at 22 °C. The pump uses outside air at −10 °C as its low-temperature reservoir and runs on electricity that costs 2.69¢/MJ. How much does it cost to keep your house at 22 °C from 8 a.m. to 10 p.m.? ••

77. The thermal gradient of Earth's crust is approximately 25 K/km at locations away from tectonic plate boundaries. Suppose you want to use this thermal gradient to run a heat engine that operates on a Carnot cycle at a location where the base of the crust is 30 km below Earth's surface and the surface temperature is 28 °C. (*a*) What is the maximum efficiency the engine can achieve at this location? (*b*) Because drilling 30 km into the crust presents great technological problems, you will settle for an efficiency of 0.50. How deep into the crust must your high-temperature reservoir be in order for the engine to have an efficiency of 0.50? ••

78. The average daytime temperature at the surface of the Moon is 380 K, and the temperature of the lunar core is estimated at 1000 K. Suppose you want to run a heat engine that operates on a Carnot cycle to use this temperature difference to run a 2.50-MW power station at some future lunar colony. (*a*) At what rate does this engine extract thermal energy from the lunar core? (*b*) At what rate does the process change the entropy of the core system? (*c*) What is the thermal efficiency of this engine? ••

79. A heat engine that operates on a Carnot cycle has efficiency η. When the engine is run in reverse as a heat pump, what are, in terms of η, (*a*) the coefficient of performance of cooling when the pump is used for cooling and (*b*) the coefficient of performance of heating when the pump is used for heating? (*c*) For each case, check your answer to see whether the range of the coefficients of performance is physically reasonable. ••

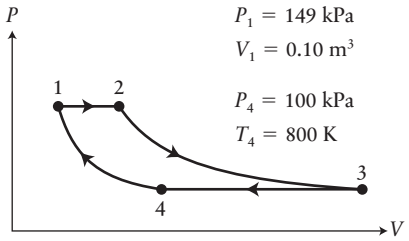
80. A heat engine that operates on a Carnot cycle is used to pump crude oil at a rate of 25,000 barrels/day from a well that is 1.70 km deep. The mass density of the oil is 850 kg/m^3, the oil is under minimal pressure and so has no initial speed, and it has negligible speed as it reaches the surface. The engine operates between 50 °F and 250 °F, and each barrel has a volume of 0.159 m^3. At what rate must thermal energy be supplied to this heat engine? ••

81. A heat engine that operates on a Carnot cycle between 150 °C and 10 °C is used to accelerate a flywheel from rest to a rotational speed of 8.50 s^{-1}. The 1500-kg flywheel is a solid uniform cylinder that rotates about its long central axis without appreciable friction. Its diameter is 1.00 m, and it is 750 mm long. How much thermal energy must be transferred into this engine to accelerate the flywheel to the desired rotational speed, and how much of this thermal energy is wasted in the process? ••

82. A 950-kg car uses an engine that operates on a Brayton cycle for which the pressure ratio is 8 and the working substance is helium (He) gas. During each cycle, 25% of the work done by the engine is used to overcome friction. If the engine must accelerate the car from rest to 30 m/s, how much energy must be transferred thermally to the engine? ••

83. The Brayton cycle represented in Figure P21.83 operates with 2.00 mol of working substance. How many degrees of freedom does each particle of the working substance have? ●●●

Figure P21.83



84. You have taken a job as a patent examiner, and your first assignment is to evaluate a "zero-point energy perpetual motion machine." You are skeptical, but after the applicant gives it an initial energy of 1000 J, the machine starts operating and does not immediately run down. Hours later, tired of waiting for the machine to stop, you use an infrared sensor and determine that the machine is losing thermal energy at an average rate of 200 μJ/s. You would like to examine other devices while waiting for the machine to stop, but you do not want to miss its final few minutes of operation, and you are also concerned about its effect on the entropy of the environment. ●●● CR

85. Working for an automotive journal, you are reviewing two proposed new models that have identical frame shapes and identical masses. Car 1 runs on a Brayton cycle that does enough work to accelerate the car from 0 to 100 km/h in 11.5 s when internal friction, air resistance, and other retarding factors are left out of the equations. When air resistance is taken into account, the car's top speed is 105 km/h. In the cycle, the isobaric cooling is accomplished by exposing the working substance to normal atmospheric pressure and 350 K, the temperature under the car's hood. When the working substance is compressed, the pressure increases by a factor of 5.00, and the temperature of the working substance after the isobaric combustion is 700 K. At the compressor, the engine takes in 1.00 mol of the working substance (which the designer has modeled as having four degrees of freedom) and the operating speed is 8.00 cycles/s. Car 2 runs on a Carnot cycle that has the same minimum and maximum temperatures of 350 K and 700 K, the same fuel intake, and the same piston volume. The frame shapes and masses of the cars are nearly identical. The only difference (other than the cycle itself) is that car 2 runs at only 1.00 cycle per second. ●●● CR

86. You work for a toy company that wants to produce a bathtub toy boat that operates on a Carnot cycle and slowly putts around the bathtub. It is assumed that the temperature of the bathwater will be about 100 °F and the air temperature will be about 70 °F. After some experimenting, you determine that your prototype model operates at about 40% of the efficiency of an ideal Carnot cycle between these two temperatures. Your engineer tells you that the boat needs to have an efficiency of 0.045 in order to move around well in the tub. You begin to consider simple design modifications in an attempt to reach this goal. ●●● CR



22

Electric Interactions

22.1 Static electricity
22.2 Electrical charge
22.3 Mobility of charge carriers
22.4 Charge polarization

22.5 Coulomb's law
22.6 Forces exerted by distributions of charge carriers

CONCEPTS

QUANTITATIVE TOOLS

Electricity is a familiar term—outlets, batteries, light bulbs, computers all involve electricity. It is no understatement to say that modern life depends on electricity, but what exactly *is* electricity? We all know what electricity does, but it's not that easy to explain what electricity *is*.

Electricity manifests itself in many ways: from the sparks that fly when you scuff your feet across a carpet on a dry winter day to the electricity we use in our homes to the transmission of radio and television programs. Even the attraction between magnets has to do with electricity. In this chapter, we begin our treatment of electricity with a discussion of static electricity.

22.1 Static electricity

When you tear off some plastic wrap from its roll, the wrap is attracted to anything that gets close: your hand, the countertop, a dish. This interaction between the plastic wrap and other objects doesn't have to involve any physical contact. For example, you can feel the presence of a piece of freshly torn-off plastic wrap with your cheek or the back of your hand even when your face or hand is held some distance away from the piece. You may have experienced many similar interactions: Styrofoam peanuts are attracted to your arms when you unpack a box full of them (**Figure 22.1**). Running a comb through your hair on a dry day causes the comb to attract your hair. After rubbing a balloon against a woolen sweater, you can hold the balloon close to a wall and *see* the attraction as the balloon moves toward the wall. In all these instances, the mass of the objects is too small for the interactions to be gravitational. What, then, is this interaction?

You may never have thought of these interactions as being particularly strong, but consider this: If you rub a comb through your hair and then pass the comb over some small bits of paper, the bits of paper jump up to your comb and stick to it. In other words, the bits of paper accelerate upward, which means the force exerted by your comb on them must be *greater* than the gravitational force exerted on them by Earth!

Figure 22.1 Styrofoam peanuts cling to a person's arms because of static electricity.



Now try this: Quickly pull a 20-cm strip of transparent tape* out of a dispenser and suspend it from the edge of a table (just be sure the table is not metal). Notice how the tape is attracted to anything brought nearby. It might even take some practice to prevent the tape from curling up and sticking to the underside of the table or to your hand. Bring a few objects near the suspended tape and notice the attractive interaction between them.† Go ahead—experiment!

22.1 Suspend a freshly pulled piece of transparent tape from the edge of your desk. (*a*) What happens when you hold a battery near the tape? Does it matter whether you point the + side or the − side of the battery toward the tape? Does a spent battery yield a different result? Does a wooden object yield a different result? (*b*) What happens when you hold a strip of freshly pulled tape near the power cord of a lamp? Does it make any difference if the lamp is on or off?

All these interactions involving static electricity are examples of **electric interactions.** The experiment you just did tells you there is no obvious connection between electric interactions and the electricity we think of as "flowing" in electric circuits and batteries. In Chapter 31 we shall see, however, that the two are connected.

Objects that participate in electric interactions exert an **electric force** on each other. The electric force is a field force (see Section 8.3): Objects exerting electric forces on each other need not be physically touching. As you may have noticed from the interaction between the strips of tape and various nearby objects, the magnitude of the electric force depends on distance: It decreases as you increase the separation.

22.2 Suspend a freshly pulled strip of transparent tape from the edge of your desk. (*a*) Pull a second strip of tape out of the dispenser and hold it near the first strip. What do you notice? (*b*) Does it matter which sides of the strips you orient toward each other?

As Checkpoint 22.2 makes clear, not all electric interactions are attractive. Even if you increase the mass of the strips by suspending paper clips from them, the repulsion between the strips is great enough to keep the paper clips apart (**Figure 22.2**). Now place your hand between two repelling strips and notice how both strips fly toward your hand! Then run each strip of tape several times between your fingers and notice how the electric interaction diminishes or even disappears.

*For best results, use the type called "magic" tape.

†If you find something that *repels* the tape, wipe the entire surface of the object with your hand and see if it still repels—it shouldn't. Mystified? Hang on! We'll soon be able to resolve your questions.

CONCEPTS

Figure 22.2 Strips of tape just pulled out of a dispenser repel each other. The repulsive force is great enough to keep the strips apart even when they are weighted down by paper clips.



22.3 Suspend two freshly pulled 20-cm strips of transparent tape from the edge of your desk. Cut two 20-cm strips of paper, making each strip the same width as the tape, and investigate the interactions between the paper strips and the tape by bringing them near each other. Which of the following combinations display an electric interaction: paper-paper, tape-paper, tape-tape?

22.2 Electrical charge

As we saw in the preceding section, electric interactions are sometimes attractive and sometimes repulsive. In addition, the experiment you performed in Checkpoint 22.3 demonstrates that paper strips, which do not interact electrically with each other, do interact electrically with transparent tape. What causes these interactions? To answer this question, we need to carry out a systematic sequence of experiments.

Figure 22.4

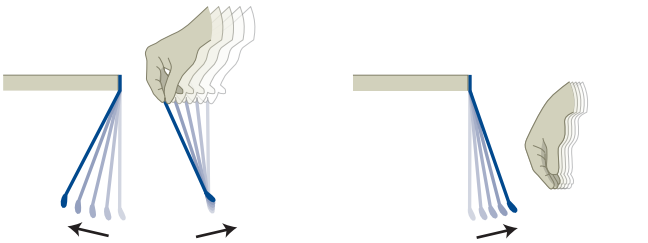


Figure 22.3 illustrates a simple procedure for reproducibly creating strips of tape that interact electrically. A suspended strip created according to this procedure interacts in the following ways: It repels another strip created in the same manner, and it attracts any other object that does not itself interact electrically with other objects (**Figure 22.4**).

Let us call the attribute responsible for the electric interaction **electrical charge,** or simply **charge.** Saying that something carries an electrical charge is just another way of saying that that object interacts electrically with other objects that carry electrical charge. Freshly pulled strips of tape carry electrical charge, and two such strips interact because each possesses an electrical charge, just as your body and Earth interact because each possesses mass. The general term for any microscopic object that carries an electrical charge, such as an electron or a proton, is **charge carrier.**

It is not immediately clear what attributes to assign to objects that do not interact electrically with each other but do interact with a charged tape strip—a strip of paper, your hand, an eraser, you name it. All we know for now is that the interaction between these objects and a charged tape is attractive rather than repulsive.

The electric charge on an object is not a permanent property; if you let a charged strip of tape hang for a while, it loses its ability to interact electrically. In other

CONCEPTS

Figure 22.3 Procedure for making strips of transparent tape that interact electrically. The purpose of the foundation strip is simply to provide a standard surface.

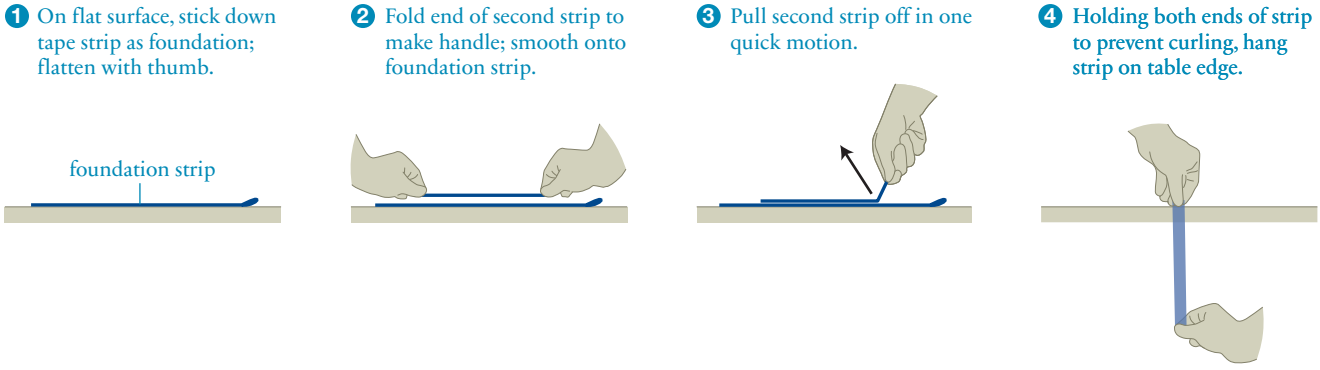
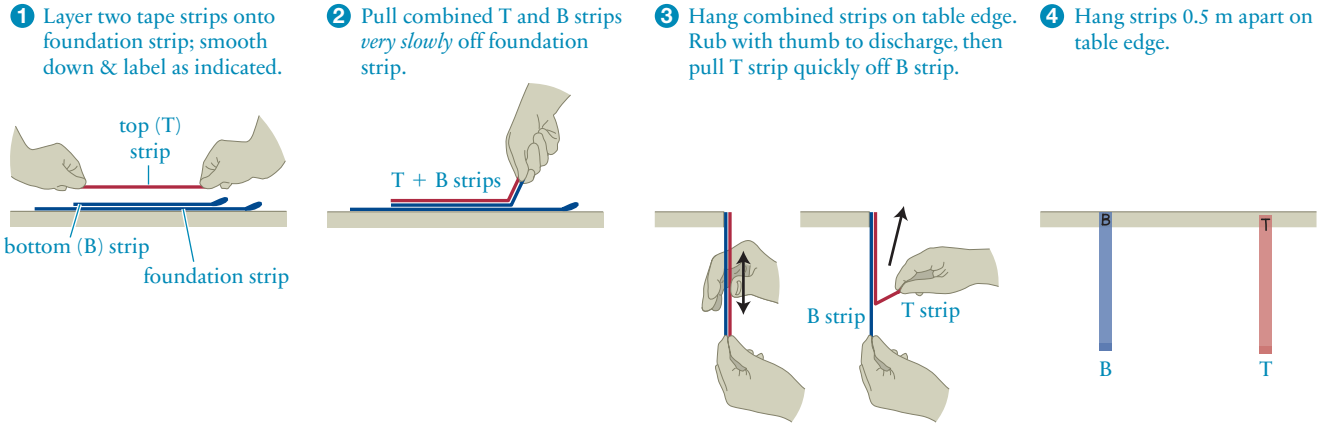


Figure 22.5 Procedure for making strips of transparent tape that carry opposite charges.



words, the strip is no longer charged—it is *discharged*. Depending on the humidity of the air, the discharging can take minutes or hours, but you can speed up the discharging by rubbing your fingers a few times over the entire length of a suspended charged strip of tape.* (The rubbing allows the charge to "leak away" from the tape by distributing itself over your body.)

22.4 (*a*) Prepare a charged strip of transparent tape as described in Figure 22.3 and then suspend the strip from the edge of your desk. Verify that the tape interacts as you would expect with your hand, with a strip of paper, and with another charged strip of tape. (*b*) Rub your fingers along the hanging strip to remove all the charge from it, and then verify that it no longer interacts with your hand. If it does interact, rub again until it no longer interacts. (*c*) Predict and then verify experimentally how the uncharged suspended strip interacts with a strip of paper and with a charged strip of tape.

CONCEPTS

To restore the charge on a discharged strip, stick the strip on top of the foundation strip from which you pulled it off (step 1 in Figure 22.3), smooth it out, and then quickly pull it off again. You can recharge a strip quite a few times before it loses its adhesive properties. Once the tape does lose its adhesiveness, however, recharging it becomes impossible. It is generally a good idea to rub your finger over the foundation strip before you reuse it to make sure that it, too, is uncharged.

22.5 Recharge the discharged strip from Checkpoint 22.4 and verify that it interacts as before with your hand, with a strip of paper, and with another charged strip of tape.

A discharged tape strip interacts in the same way as objects that carry no charge. Such objects are said to be electrically **neutral.** They do not interact electrically with other neutral objects, but they do interact electrically with charged objects. We shall examine this surprising fact in more detail in Section 22.4.

Where does the electrical charge on a charged tape strip come from? Is charge *created* when two strips are pulled apart as in Figure 22.3? This is something we can check by sticking two strips of tape together, rubbing with our fingers to remove all charge from the combination, and then quickly separating the two strips (**Figure 22.5**).

22.6 Follow the procedure illustrated in Figure 22.5 to separate a pair of charged strips. (*a*) How does strip B interact with a neutral object? How does strip T interact with a neutral object? (*b*) Create a third charged strip and see how it interacts with strip B and with strip T. (*c*) Is strip T charged? (*d*) Is strip B charged? (*e*) Check what happens to the interactions with B and T strips when you discharge a B or a T strip by rubbing your fingers along its length.

As Checkpoint 22.6 shows, separating an uncharged pair of strips produces two charged ones, but the behavior of strip B is different from that of the other strips we have encountered so far!

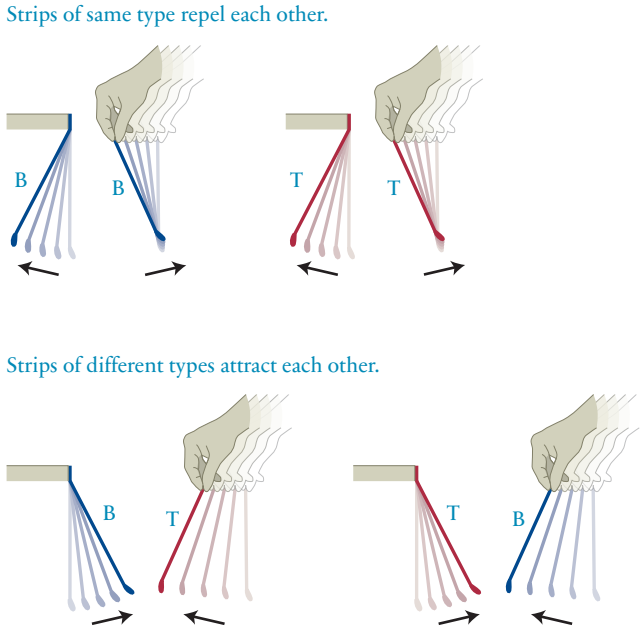
22.7 Make two charged pairs of strips (B and T) following the procedure illustrated in Figure 22.5. Investigate the interaction of B with T, T with T, and B with B.

The interactions between the B and T strips are illustrated in **Figure 22.6**: Strips of the same type repel each other, while strips of different types attract each other. This series of experiments leads us to conclude that there are two types of charge on the tapes, one type on B strips and another type on T strips. Strips that carry the same type of charge, called *like charges*, exert repulsive forces on each other; strips that carry different types of charge, called *opposite charges*, exert attractive forces on each other.

Having determined that two types of electrical charge exist, a logical next question is: Are there even more types?

*If rubbing your fingers along the tape doesn't do the job, try wetting them before rubbing them over the tape.

Figure 22.6 Interactions of B and T charged strips.



22.8 (*a*) Prepare one charged strip of tape according to Figure 22.3 and hang it from the edge of your desk. Hang a narrow strip of paper from the desk edge also, about 0.5 m away from the tape strip. Pass a *plastic* comb six times quickly through your hair and then show that the comb is charged. Be sure to use a plastic comb; combs made from other materials do not acquire a charge when passed through hair. The cheapest type of comb usually works best. (*b*) Make a pair of oppositely charged B and T strips (Figure 22.5) and investigate how they interact with a charged comb. (*c*) Does your comb behave like a B strip, a T strip, or neither?

Experiments show that *any* charged object—obtained by rubbing objects together or otherwise—always attracts either a B strip or a T strip and repels the other. No one has ever found a charged object that repels or attracts *both* types of strips. In other words:

There are two and only two types of charge. Objects that carry like charges repel each other; objects that carry opposite charges attract each other.

The two types of charge never appear independently of each other: Whenever two neutral objects are either rubbed together and then separated or, if an adhesive surface is involved, stuck together and then separated and one of them acquires a charge of one type, the other object always acquires a charge of the other type. The generation of opposite charges is obvious when you separate a neutral pair of tape strips. When you pass a comb through your hair, the comb acquires a charge of one type and your hair acquires a charge of the other type. On a dry day, you may have noticed that some hair strands stand up away from your head. Each charged strand is being repelled by the other charged strands, and so they are all getting as far away from one another as possible.

It can be shown that when two tape strips are separated, the forces exerted by the B strip and the T strip on a third charged strip are equal in magnitude, although one is attractive and the other repulsive. Furthermore, when the B and T strips are recombined, the combination is neutral again. These observations suggest that after you rub and then separate a pair of objects, the objects carry equal amounts of opposite charge. Combining these equal amounts of opposite charge produces zero charge. These observations indicate that all neutral matter contains equal amounts of positive and negative charge. The two types of charge are called **positive** and **negative charges.** The definition of negative charge is as follows:*

Negative charge is the type of charge acquired by a plastic comb that has been passed through hair a few times.

22.9 Does the B strip you created in Checkpoint 22.8 carry a positive charge or a negative charge?

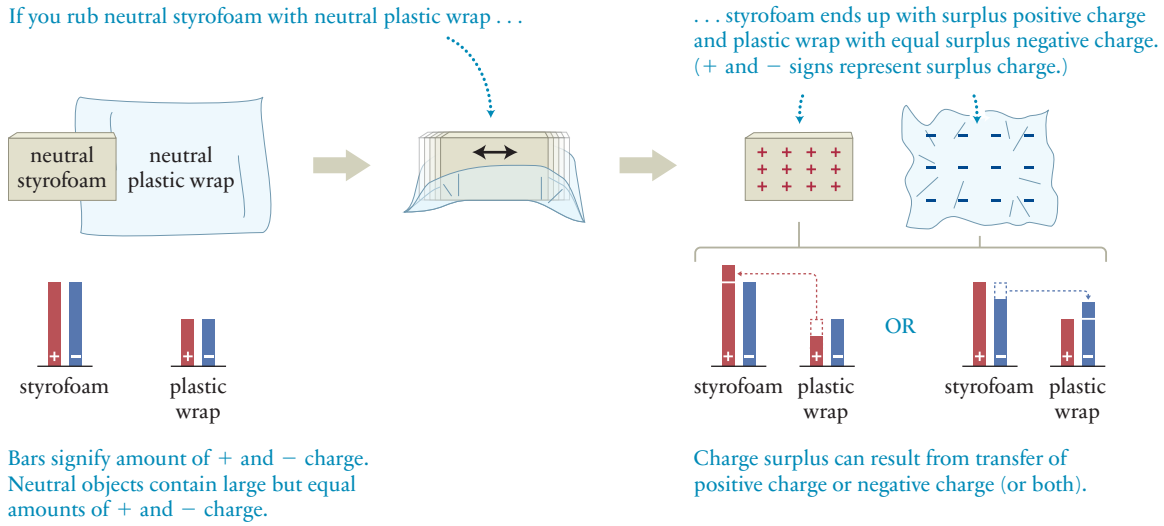
When two neutral objects touch, some charge can be transferred from one object to the other, with the result that one object ends up with a surplus of one type of charge and the other object ends up with an equal surplus of the other type of charge. For example, when a neutral piece of styrofoam is rubbed with a neutral piece of plastic wrap, the styrofoam acquires a positive charge (meaning it contains more positive than negative charge) and the plastic wrap acquires a negative charge (it has a surplus of negative charge). Without further information, however, we cannot tell whether positive charge has been transferred from the wrap to the styrofoam, or negative charge has been transferred from the styrofoam to the plastic wrap, or a combination of these two. (See **Figure 22.7** on the next page.) Summarizing:

All neutral matter contains equal amounts of positive and negative charge; charged objects contain unequal amounts of positive and negative charge.

In illustrations, surplus charge is represented by plus or minus signs. Keep in mind, however, that these signs never represent the only type of charge in an object. The plus signs on the positively charged styrofoam in Figure 22.7, for example, mean only that the styrofoam

*Historically, negative charge was (arbitrarily) defined by Benjamin Franklin (1706–1790) as the charge acquired by a rubber rod rubbed with cat fur. Because plastic combs and hair are more easily accessible than rubber rods and cat fur, the definition of negative charge given here is more convenient.

Figure 22.7 Rubbing neutral styrofoam with neutral plastic wrap leaves the two objects with equal charges of opposite types.



contains more positive than negative charge, either because some of its negative charge has been removed or because some positive charge has been added. In addition to the 12 positive charge carriers shown in Figure 22.7, the styrofoam contains millions and millions of positive charge carriers paired with millions and millions of negative charge carriers. A drawing such as Figure 22.7 shows only *unpaired* charge carriers (usually referred to as *surplus charge*).

As our observations in Figure 22.6 show, oppositely charged B and T strips attract each other. The interaction between positive and negative charge tends to bring positive and negative charge carriers as close together as possible. Because combining equal amounts of positive and negative charge results in zero charge, we can say that charge carriers always tend to arrange themselves in such a way as to produce uncharged objects—indeed, all matter around us tends to be neutral.

CONCEPTS

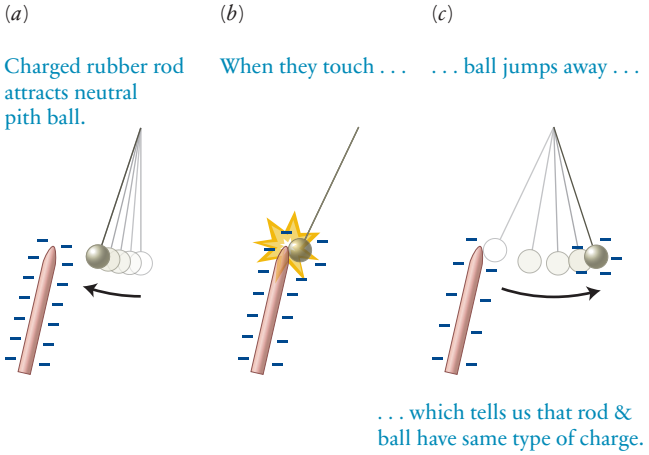
22.10 Imagine having a collection of charged marbles that retain their charge even when they touch other objects. Red marbles are positively charged, and blue marbles are negatively charged. (*a*) What happens if you place a bunch of red marbles close together on a flat horizontal surface? (*b*) What happens if you do the same with a bunch of blue marbles? (*c*) What happens if you do the same with an equal mixture of red and blue marbles? (*d*) What happens in part *c* if you have a few more red marbles than blue ones? (*e*) As a whole, is the collection of marbles in part *d* positively charged, negatively charged, or neither?

22.3 Mobility of charge carriers

To gain a better understanding of electrical charge, many additional experiments are required, most of which require items not easily found at home. A rubber rod rubbed with a piece of cat fur acquires—by Benjamin Franklin's original definition—a negative charge (and the fur acquires a positive charge). A glass rod rubbed with silk acquires a positive charge (and the silk a negative charge). Other materials also acquire a charge upon contact or rubbing, but these two combinations of rubber/fur and glass/silk provide the most convenient means of generating relatively large amounts of charge.

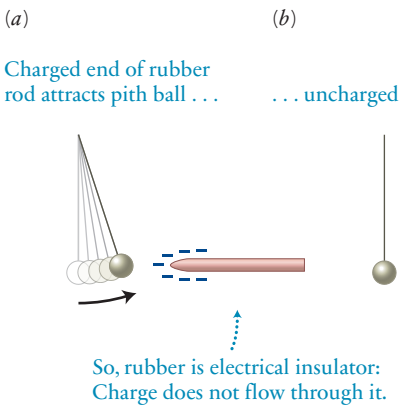
Interesting things happen when a charged rubber rod is brought into contact with an uncharged pith ball.* As the rod is brought near the ball, the ball moves toward the rod because of the attraction between the charged rod and the neutral ball (**Figure 22.8*a***). As the ball touches the rod, however (Figure 22.8*b*), the crackling sounds of tiny sparks may be heard. The ball suddenly jumps away from the rod (Figure 22.8*c*), indicating that the interaction between rod and ball has become repulsive. This repulsive interaction indicates that the ball has acquired the same type of charge as the rod (negative).

Figure 22.8 A charged rubber rod can transfer charge to a neutral pith ball.



Pith is the soft, lightweight, spongelike material that makes up the interior of the stems of flowering plants.

Figure 22.9 A rubber rod is an example of an electrical insulator.



In other words, some of the surplus negative charge on the rod has been transferred to the ball.

Charge can be transferred from one object to another by bringing the two into contact.

We can use this phenomenon to investigate the electrical behavior of different kinds of materials. For example, if we transfer some charge to one end of an uncharged rubber rod and then extend the charged end toward an uncharged pith ball, the two interact electrically, as shown in **Figure 22.9*a***. If we hold the *uncharged* end near the pith ball, as in Figure 22.9*b*, however, no interaction occurs. This tells us that the charge does not flow from one end of the rubber rod to the other; instead, it remains near the spot where it has been deposited. Any material in which charge doesn't flow (or moves only with great difficulty) is called an **electrical insulator.**

Electrical insulators are materials through which charge carriers cannot flow easily. Any charge transferred to an insulator remains near the spot at which it was deposited.

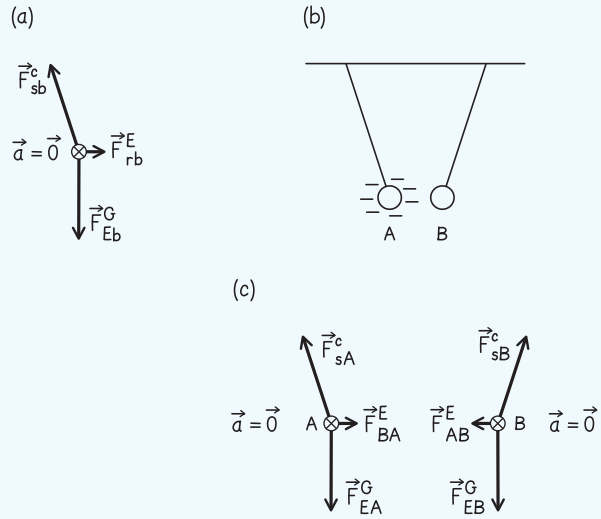
Glass, rubber, wood, and plastic are examples of electrical insulators. Air, particularly dry air, is also an insulator, although the presence of large amounts of charge can cause charge carriers to "jump" from one object to another, causing sparks.

EXERCISE 22.1 Electric forces

(*a*) Draw a free-body diagram for the pith ball in Figure 22.9*a*. (*b*) Two identical neutral pith balls A and B are suspended side by side from two vertical strings. After some charge is transferred from a charged rubber rod to A, A and B interact. (B remains neutral because the two balls never come into contact with each other.) Sketch the orientation of A and B after the charge has been transferred to A. (*c*) Draw a free-body diagram for each ball.

SOLUTION (*a*) The ball is subject to three forces: a gravitational force, a contact force exerted by the string, and an attractive electric force exerted by the charged particles in the rod. This last force is directed horizontally toward the rod. Because the ball is at rest, I know that the vector sum of these three forces is zero. So the horizontal component of the force exerted by the string on the ball must be equal in magnitude to the electric force exerted by the rod on the ball, and the vertical component of the force exerted by the string on the ball must be equal in magnitude to the gravitational force exerted by Earth on the ball (**Figure 22.10*a***).

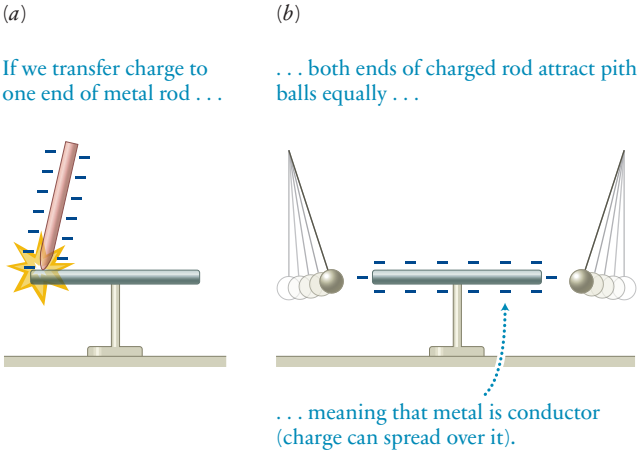
Figure 22.10



(*b*) As we saw in Section 22.2, a neutral object interacts electrically with a charged object. The electric force exerted by A on B and that exerted by B on A form an interaction pair and so their magnitudes are equal. Because the masses of the pith balls are the same, each is pulled in by the same distance. Thus my sketch is as shown in Figure 22.10*b*. (*c*) See Figure 22.10*c*.

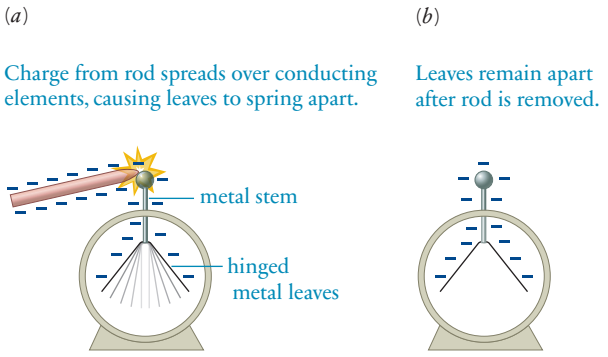
In **Figure 22.11**, a charged rubber rod is brought into contact with an uncharged metal rod supported on an electrically insulating stand. Once the charged rod has touched the metal rod, *all* points on the surface of the

Figure 22.11 A metal rod is an electrical conductor.



CONCEPTS

Figure 22.12 An electroscope depends on electrical conduction.



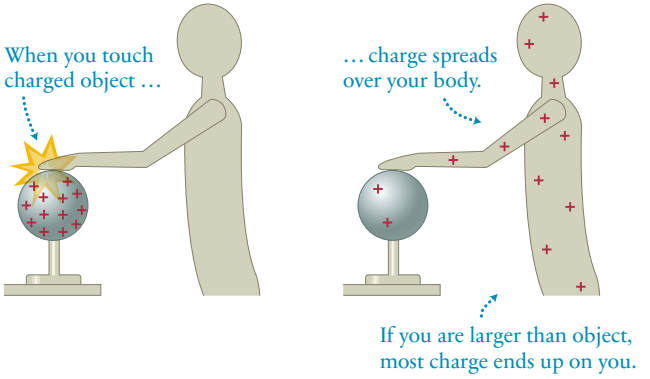
metal rod interact electrically with other objects, indicating that the charge spreads out over the metal rod. The tendency of charge to spread out over metal objects can be demonstrated with an *electroscope* (**Figure 22.12*a***). Two strips made of metal foil are suspended from a small metal rod in an electrically insulating enclosure; the rod is connected to a metal ball on top of the enclosure. When the metal ball is charged by an exterior source, the strips move away from each other. The explanation for this movement is that the added charge quickly moves from the metal ball through the metal rod and onto the two metal strips. Once the strips carry the same type of charge, they repel each other (Figure 22.12*b*).

Another demonstration of the free motion of charges through metals is shown in **Figure 22.13**: When a long wire is used to connect a charged metal sphere to an uncharged metal sphere, charged particles flow from the charged sphere to the uncharged one. Because wires are made of metal, this experiment shows that, in contrast to what happens with electrical insulators, charge moves easily through a metal and across a metal-to-metal contact. Materials through which charge carriers can flow easily are called electrical **conductors,** and the flow of charge carriers through conductors is called **conduction.**

Electrical conductors are materials through which charge carriers can flow easily. Any charge transferred to a conductor spreads out over the conductor and over any other conductors in contact with it.

Metals are the only solid materials that are conductors at room temperature. (As noted earlier, glass, plastic, and most other solids are electrical insulators.) Although charge does not flow easily through pure water, minute amounts of impurities turn water into a good conductor. Because most water contains some impurities, water is therefore usually considered a conductor.

Figure 22.14 Because of its water content, the human body is a conductor.

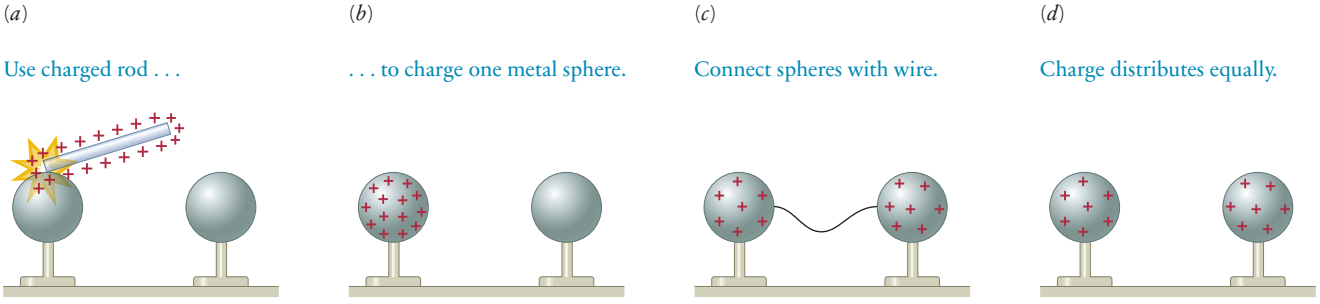


Except for the outer layer of soil, Earth is also a good conductor. Consequently, when a charged, conducting object is connected to Earth by a wire, a process called **grounding,** charge carriers can flow between Earth ("ground") and the object. Because Earth is so large, it can supply or absorb a nearly unlimited number of charge carriers. In the absence of other nearby electrical influences, the grounded object is left with no surplus of either type of charge.

Because of its high water content, the interior of the human body is a conductor. Consequently, any time you touch a charged object, as in **Figure 22.14**, some of the charge moves into you—you act like a grounding agent just the way Earth does. As long as you keep touching the object, charge flows into your body, reducing the charge on the object (Figure 22.14). If the charge on the object is large, the charge that accumulates on your hair makes your hair stand up and separate as far as possible, like the leaves of an electroscope (**Figure 22.15**).

22.11 (*a*) Why is it impossible to charge a metal rod held in your hand by rubbing the rod with other materials? (*b*) Why can you charge a rubber rod even when you hold it in your hand?

Figure 22.13 A conducting wire distributes charge between two conducting spheres.



CONCEPTS

Figure 22.15 Charge spreads over the human body, so a large charge will cause your hairs to repel one another and stand on end.



Is charge some sort of fluid that flows from one object to another, or is it composed of small particles that can be peeled off or stuck onto objects? To answer this question, we must look at the atomic structure of matter. All matter on Earth consists of atoms (see Section 1.3), the structure of which is schematically illustrated in **Figure 22.16**. Nearly all the atom's mass is concentrated in the extremely small nucleus at the center. The nucleus is composed of protons and neutrons. The region surrounding the nucleus, representing most of the atom's volume, is a cloud of electrons.

Experiments show that electrons have a negative electrical charge—they are repelled by a charged comb* and by other electrons. Protons carry a positive charge, and neutrons carry no charge. The protons and neutrons in the nucleus are held together tightly by the strong interaction (see Chapter 7), which is great enough to overcome the electrical repulsion between the positively charged protons. The electrical attraction between the positively charged nucleus and the negatively charged electrons is responsible for keeping the electrons bound to the nucleus. The electron cloud does not collapse on the nucleus because of additional constraints imposed on the electrons by the laws of quantum mechanics.

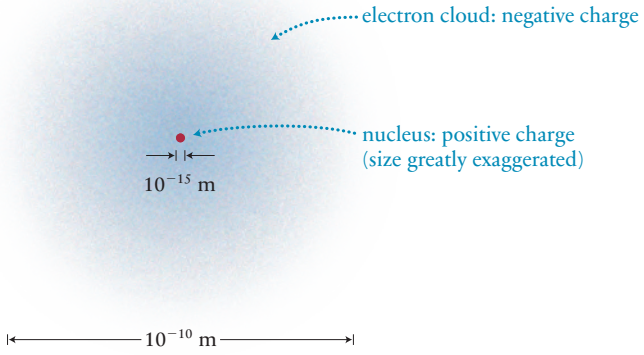
Charge is an inherent property of the electron, which means it is impossible to remove the charge from an electron—there is no such thing as a discharged electron. Experiments show that, excluding quarks (Section 7.6),

All electrical charge comes in whole-number multiples of the electrical charge on the electron.

For this reason, the magnitude of the charge on the electron, designated by the letter e, is called the **elementary charge.**

*Recall from our discussion of positive and negative charge carriers in Section 22.2 that a plastic comb carries a negative charge.

Figure 22.16 Structure of the atom (not to scale).



Every atom contains equal numbers of electrons and protons. Because atoms are neutral, the fact that they contain equal numbers of electrons and protons tells us that the magnitude of the positive charge on the proton is also e. The charge on an electron is $-e$, and that on a proton $+e$. As with electrons, the charge cannot be removed from a proton.

Given that macroscopic objects contain an immense number of atoms and that each atom can contain dozens of electrons and protons, we see that ordinary objects contain an immense number of positively charged protons, exactly balanced by an equal number of negatively charged electrons. A surplus of just a minute fraction of these numbers is sufficient to give rise to a noticeable macroscopic charge. For example, when you pull apart two strips of transparent tape, the separation causes a surplus of less than one in a trillion (10^{12}) electrons. (Because there are about 10^{22} electrons in the strip, that fraction represents some 10^{10}, or ten billion electrons.)

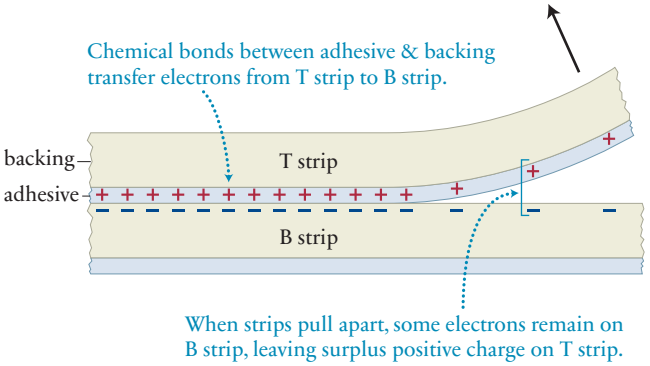
When two atoms are brought close together, they may form a chemical bond by transferring or sharing one or more electrons from one atom to the other. When an electron transfer takes place, both atoms contain unequal numbers of electrons and protons and are now called **ions** instead of atoms. One of the two ions has gained one or more electrons, meaning it contains more electrons than protons and therefore carries a negative charge. The other ion, the one that lost electrons, contains more protons than electrons and so carries a positive charge.

Ions in solids move very little, but ions in liquids can move freely. For instance, in table salt, a compound made of pairs of sodium (Na^+) and chloride (Cl^-) ions, the charged ions hardly move at all, meaning that solid table salt is an electrical insulator. Dissolve table salt in water, however, and the solution contains large quantities of positively charged sodium ions and negatively charged chloride ions. Because these ions can move freely, the solution is an electrical conductor.

Some solids are made not of paired ions the way sodium chloride is but rather of individual atoms. In atomic solids that are electrical insulators, the electrons

CONCEPTS

Figure 22.17 How strips of tape can acquire opposite charges when pulled apart.



in the atoms are unable to move because each electron is bound to a specific atom. Diamond (made of the element carbon) and glass are two familiar examples. Metals are also atomic solids rather than ionic solids, but in metals, each atom gives up one or more electrons to a shared "gas" of electrons that spreads throughout the volume of the metal. The metal as a whole is still neutral: The negative charge of the electron gas is exactly balanced by the positive charge of the ions. The electrons in the gas are called *free electrons* because they can move freely inside the metal; these electrons are responsible for the easy flow of charge through a metal.

Nearly all electrical phenomena are due to the transfer of electrons—and therefore charge—from one atom to another. For example, when the sticky side of one strip of transparent tape is applied on top of the nonsticky side of a second strip, atoms in the adhesive from the top strip form chemical bonds with atoms in the nonsticky surface of the bottom strip by transferring electrons, as shown in **Figure 22.17**. These bonds are responsible for the adhesion of one strip to the other. When the strips are pulled apart quickly, the bonds are broken, but not all electrons manage to get back to the top strip. The bottom strip thus ends up with a surplus of electrons, making it negatively charged, and the top one with a deficit of electrons, making it positively charged.*

As we saw in Section 10.4, friction between two surfaces also involves the breaking of chemical bonds. As with the separation of the tape strips, this bond breaking sometimes leaves surplus charge on the surfaces. When you touch a piece of plastic food wrap to a piece of styrofoam, for instance, chemical bonds form between atoms on the two surfaces. In these bonds, electrons from the styrofoam move to the wrap. If you then rub the surfaces against each other, these bonds are broken, and some of the electrons originally on the styrofoam stay on the wrap. If the breaking of the chemical bonds occurs slowly, the electrons migrate back and no surplus charge builds up. For that reason it is necessary to rub vigorously or to separate strips of tape quickly. The key point is:

*Depending on the type of adhesive and the material of the backing, the transfer of electrons can also be in the other direction.

Any two dissimilar materials become charged when brought into contact with each other. When they are separated rapidly, small amounts of opposite charge may be left behind on each material.

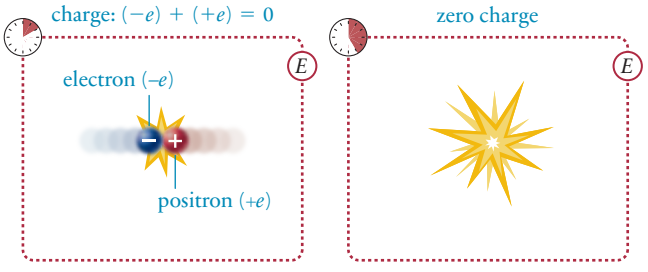
Because charging by breaking of chemical bonds is due to a transfer of charge, we now see that for every surplus of negative charge that appears in one place, an equal surplus of positive charge appears somewhere else. After the two strips in Figure 22.17 are separated, the sum of the positive charge on the T strip and the negative charge on the B strip is still zero. No creation or destruction of charge is involved, suggesting that electrical charge—like momentum, energy, and angular momentum—is a conserved quantity. The principle of **conservation of charge** states:

Electrical charge can be created or destroyed only in identical positive-negative pairs such that the charge of a closed system always remains constant.

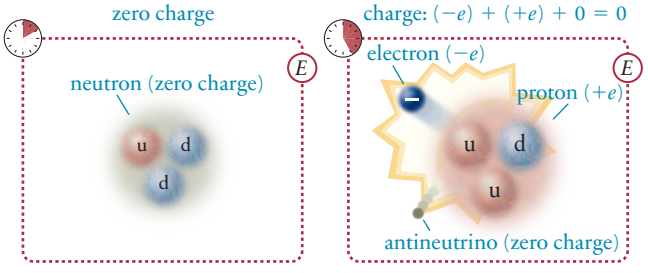
No process has ever been found to violate this principle. Even when charged subatomic particles, such as electrons and protons, are created or destroyed—a process that can be observed in high-energy particle accelerators—charge is conserved. For example, when an electron (charge $-e$) collides with a subatomic particle called the *positron* (charge $+e$), both particles are destroyed, leaving nothing but a flash of highly energetic radiation (**Figure 22.18***a*). The charge of the electron-positron system before the collision is $(-e) + (+e) = 0$, and it is still zero after the collision. Likewise, in a process called beta decay, when a free neutron (carrying

Figure 22.18 Because charge is conserved, the charge of a closed system does not change even when particles are created or destroyed.

(*a*) Electron-positron annihilation



(*b*) Decay of a free neutron



CONCEPTS

zero charge and made up of two down quarks and one up quark) decays into a proton (charge $+e$, one down and two up quarks), an electron (charge $-e$), and a neutral subatomic particle called the antineutrino (zero charge), the charge of the system comprising the neutron and the particles into which it decays remains zero (Figure 22.18*b* and Section 7.6).

22.12 When two objects made of the same material are rubbed together, friction occurs but neither material acquires surplus charge. Why?

22.4 Charge polarization

Let us now reexamine the interaction between a charged object and a neutral one. **Figure 22.19** shows the interaction between a charged rubber rod and an uncharged electroscope. With the rod far away (Figure 22.19*a*), the leaves of the electroscope hang straight down. When the rod is brought near the ball of the electroscope (Figure 22.19*b*), the leaves separate even without any contact between the rod and the electroscope. As the distance between the rod and the electroscope is increased again, the leaves drop down, showing that no charge has been transferred from the rod to the electroscope.

Why do the leaves separate even though the electroscope remains neutral? They separate because the negative charge on the rod repels the free electrons in the metallic parts of the electroscope: The free electrons are pushed as far away as possible from the rod (Figure 22.19*c*) and pile up in the leaves. This redistribution of charge is nearly instantaneous. The top of the electroscope thus ends up with a deficit of electrons—a positive charge—and the leaves end up with a surplus of electrons—a negative charge. The negative charge on the leaves is responsible for the repulsion between them. When the rod is removed, the electrons, being repelled by one another and attracted to the positive charge on the electroscope ball, immediately flow back to their normal positions, evening out the distribution of positive and negative charge.

22.13 (*a*) In Figure 22.19*b*, is the electroscope as a whole positively charged, negatively charged, or neutral? (*b*) How does the magnitude of the positive charge on the electroscope ball compare with the magnitude of the negative charge on the leaves? (*c*) Is the force exerted by the rod on the electroscope ball attractive or repulsive? Is the force exerted by the rod on the leaves attractive or repulsive? (*d*) How do you expect the magnitude of the force the rod exerts on the ball to compare with the magnitude of the force the rod exerts on the leaves?

Any separation of charge carriers in an object is called **charge polarization,** or simply **polarization,** and an object in which charge polarization occurs is said to be *polarized*. The electroscope of Figure 22.19*b*, for instance, is polarized by the nearby charged rod. In any object in which charge is polarized, there are two charged *poles*, one positive and the other negative. In the electroscope of Figure 22.19*b*, the positive pole is at the ball and the negative pole is in the foil strips.

In metals, the polarization induced by the presence of a nearby charged object is very great because the free electrons in the metal move easily in response to the presence of the charged object. Even in electrical insulators, however, where there are no free electrons moving about, a nearby charged object induces some polarization. The basic reason for the polarization of insulators is illustrated in **Figure 22.20**: In the presence of an external charge, the center of the electron cloud and the nucleus of an atom shift away from each other, causing the atom to become polarized. So, when a negatively charged comb is brought near a small piece of paper, each atom in the paper becomes polarized—the electron clouds are pushed away from the comb, and the nuclei are pulled toward the comb. If we consider the paper as consisting

Figure 22.19 In (*a*) and (*b*), a charged rod induces polarization in an electroscope. (*c*) A schematic atom-level view.

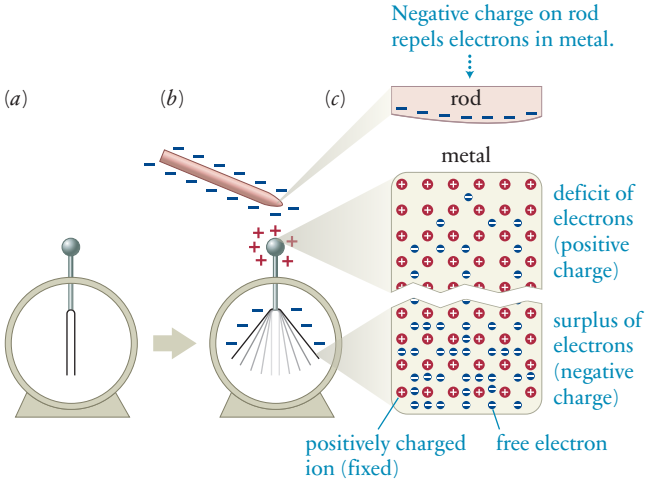
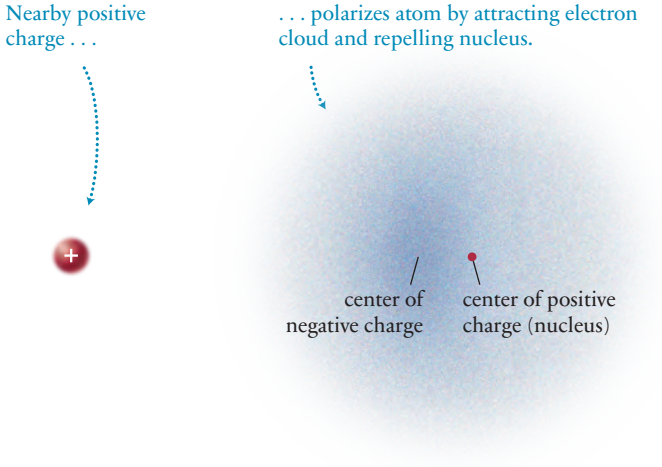
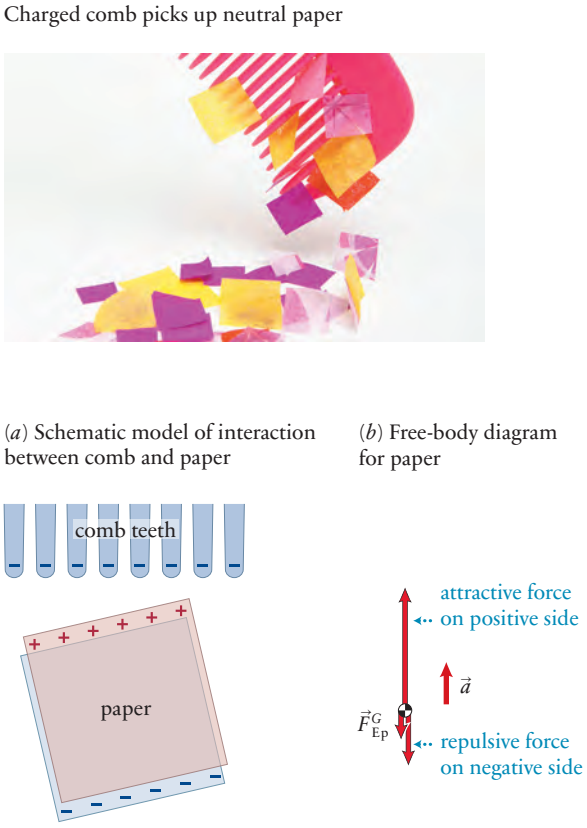


Figure 22.20 Polarization of a neutral atom.



CONCEPTS

Figure 22.21 Polarization of a neutral insulator (bits of paper) by a charged comb. In (*a*), a single bit of paper is modeled as two offset sheets with opposite charges.



of two overlapping parts that have the same shape but carry opposite charges, the positively charged part is pulled a bit toward the comb and the negatively charged part is pushed away, as shown in **Figure 22.21*a***. This leaves the central part of the paper neutral but creates a sliver of surplus positive charge on the side facing the comb and an equal amount of surplus negative charge on the opposite side, and so the paper is polarized.

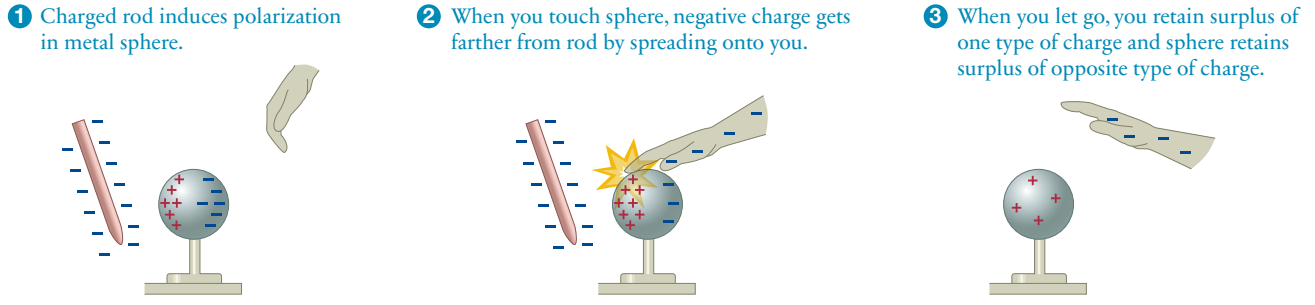
22.14 In an atom, what limits the separation between the electron cloud and the nucleus in the presence of an external charge? Why, for example, isn't the electron cloud in Figure 22.20*b* pulled all the way to the location of the external positive charge?

The polarization of atoms is responsible for the attraction between charged and neutral objects. In Figure 22.21, for example, the positively charged side of the paper is closer to the comb than the negatively charged side. Because the electric force magnitude decreases with increasing distance, the magnitude of the attractive force exerted on the positive side is greater than the magnitude of the repulsive force exerted on the negatively charged side (Figure 22.21*b*). Consequently, the vector sum of the electric forces exerted by the comb on the neutral piece of paper points toward the comb and the paper is pulled toward the comb.

22.15 (*a*) When a positively charged object is brought near a neutral piece of paper, is the vector sum of the forces exerted by the charged object on the paper attractive or repulsive? (*b*) Describe what would happen when a negatively charged comb is brought near an electroscope if protons, not electrons, were mobile in metals. (*c*) Can you deduce from the experiment illustrated in Figure 22.19 which charge carriers—electrons or protons—are mobile in metals?

Figure 22.22 illustrates how you can exploit polarization to charge neutral conducting objects. A negatively charged rod brought near a metal sphere induces polarization in the sphere. To get as far away as possible from the negative charge on the rod, the electrons in the metal of the sphere move to the right surface. Thus the surface of the sphere nearer the rod becomes the positive pole, and the surface farther from the rod becomes the negative pole. When you touch the metal, the electrons can move into your body, thereby getting even farther away from the negative charge on the rod. In essence, you have become the negative pole, while the sphere is the positive pole. If you remove your hand from the sphere, the sphere is left with a deficit of electrons (they stayed inside you!) and so carries a positive charge—it has become charged without ever touching a charged object. (Likewise, you now have a surplus of negative charge and carry a negative charge, which is spread so thin that it is hardly noticeable.) This process is called **charging by induction.**

Figure 22.22 Polarization can be exploited to charge neutral conducting objects.



CONCEPTS

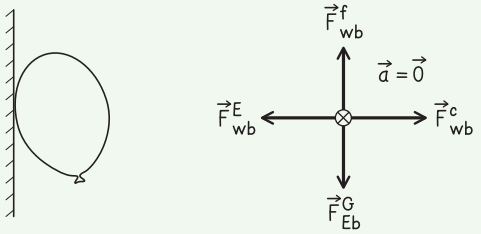
Self-quiz

1. You can use a positively charged object to charge a neutral object (*i*) by conduction (in the absence of grounding) or (*ii*) by induction (with grounding). For each process, which type of charge (positive, negative, or neither) does the neutral object acquire?
2. Because we observe two types of electric interactions, attractive and repulsive, we postulate two types of charge. Do you think there are also two types of mass? Why do you think this? Do you think there are two types of magnetic pole? Why do you think this?
3. Is the statement *A plastic comb that has been passed through the hair a few times carries a negative charge* a physical law or a definition? What are some of the differences between a law and a definition?
4. A balloon rubbed in your hair or on your clothes sticks to a wall. If you place the rubbed side of the balloon against the wall, it sticks to the wall immediately. Try this, however: After rubbing the balloon in your hair, place the side of the balloon you did not rub against the wall and notice how the balloon turns until the rubbed side is touching the wall (*a*) Draw a free-body diagram for the balloon sticking to the wall. (*b*) Given that the balloon rotates so that the rubbed area is against the wall, do you think the balloon is an electrical insulator or conductor? (*c*) Was charge created in either the balloon or the wall in order for the sticking to occur? (*d*) Is any charge transferred from the balloon to the wall? Why or why not?
5. Air can act as both an insulator and a conductor. Consider reaching for a metal doorknob after scuffing your feet over a carpet. As your hand approaches the knob, a spark jumps between your hand and the knob. Explain how air acts as both insulator and conductor in this situation.

Answers

1. (*i*) Positive. When the two objects are brought into contact with each other (a necessary condition for conduction), surplus positive charge moves from the charged object to the neutral one. Thus the neutral object acquires the same type of charge as the charged object. (*ii*) Negative. Because the objects don't touch during charging by induction, charge carriers of the same type as the charged object escape from the neutral object during grounding.
2. Mass is the quantity responsible for gravitational interactions. Because all gravitational interactions are attractive, we can assume there is only one type of mass. Magnetic poles are responsible for magnetic interactions, which can be attractive or repulsive. Therefore there must be two types of magnetic pole (called north and south).
3. Definition. A law arises from observable phenomena and is found to be true in all cases that have been tested or observed (see Section 1.1). A definition cannot be tested. There is no test that allows us to observe that the charge on a comb passed through hair is negative. All we can do is show that the charged comb behaves in a fashion similar to other objects whose charge we call negative.
4. (*a*) See **Figure 22.23**. (*b*) Insulator. Because the only portion of the balloon that is attracted to the wall is the rubbed area, we know that the charge created by the rubbing does not spread out over the balloon surface. (*c*) No. Electrical charge can never be created. The surplus charge on the balloon was transferred from your hair or clothing. (*d*) No. If change were transferred, the balloon and the wall would repel each other.

Figure 22.23



5. Scuffing your feet on the carpet transfers charge from the carpet to you. Before you get near the knob, the air insulates your charged body from the knob. As you move nearer and nearer to the knob, the magnitude of the electric force between the charge carriers in your hand and those in the knob increases until the forces are so great that the air molecules are ionized, thereby producing a conducting pathway between your hand and the knob. Now the ionized air acts as a conductor for the jumping charge.

22.5 Coulomb's law

Quantitative experiments with electrical charge are difficult to carry out because objects lose their charge and because charge carriers on objects tend to rearrange themselves in the presence of other charged objects. In the 18th century, however, the English clergyman and scientist Joseph Priestley carried out a remarkable experiment: He charged a hollow sphere and showed that no electric force was exerted on a small piece of charged cork placed inside the sphere. Remembering that Newton had proven that no gravitational force exists inside a hollow sphere (see the box "Zero gravitational force inside a spherical shell" and **Figure 22.24**, below) because the gravitational force decreases with the square of distance, Priestley proposed that the electric force, too, decreases as $1/r^2$.

In 1785, Charles Coulomb, a French physicist, provided direct evidence for an inverse-square law by measuring how the electric force between two charged spheres changes as the distance between the spheres changes. The basic apparatus for Coulomb's experiment is shown in **Figure 22.25**. A small dumbbell is suspended from a long fiber. When spheres A and B are charged, the electric force between them twists the fiber. The amount of twist is a measure of the magnitude of the force between the two spheres

Zero gravitational force inside a spherical shell

A consequence of the inverse-square dependence on distance of the gravitational force is that a uniform spherical shell exerts *no force at all* on a mass placed anywhere inside it. This result is very important in electrostatics because it provides a strong test of the law describing the electric force between charged objects.

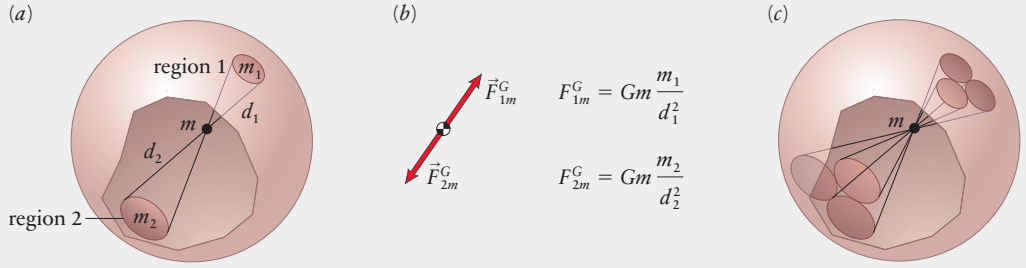
Consider the uniform spherical shell in Figure 22.24*a*. A particle of mass m is placed off-center inside the shell. To determine the force exerted by the shell on the particle, consider first the force exerted by region 1, defined as a very small region of the shell surface. Let the mass of this region be m_1 and its distance from the particle be d_1, which means the magnitude of the gravitational force exerted by this region on the particle is Gm_1m/d_1^2. Extending the cone defined by the particle and region 1 to the opposite side of the shell gives us a small region 2, of mass m_2 and at distance d_2 from the particle. The magnitude of the gravitational force exerted by this region on the particle is Gm_2m/d_2^2. If the particle is closer to region 1 than to region 2, the area of region 2 must be greater than the area of region 1 (because of how we defined region 2 as being an extension of the cone formed by region 1 and the particle), which means region 2 contains more mass: $m_2 > m_1$.

Because the mass of the shell is distributed uniformly over the shell, the mass of each of our two regions is proportional to its area: $m_1/m_2 = a_1/a_2$. Because they are marked out by similar cones, the areas of the two regions are proportional to the squares of the distances to the particle: $a_1/a_2 = d_1^2/d_2^2$, which means that

$$\frac{m_1}{m_2} = \frac{d_1^2}{d_2^2}.$$

Rearranging terms, we get $m_1/d_1^2 = m_2/d_2^2$ *regardless* of the position of the particle, and so we see that the forces cancel: The two regions exert forces of equal magnitude in opposite directions on the particle (Figure 22.24*b*). We can now apply the same arguments to other pairs of small regions on either side of the particle, each yielding equal and opposite forces on the particle (Figure 22.24*c*). The vector sum of all the forces exerted on the particle by all the small regions making up the spherical shell is thus zero.

Figure 22.24



QUANTITATIVE TOOLS

(see also Section 15.7). A similar arrangement was used a few years later by Cavendish to study gravitational interactions (see Section 13.5).

Coulomb also devised a method for systematically varying the "quantity of charge" q on a metal sphere. He found that when a charged metal sphere is brought into contact with an identical uncharged metal sphere, the final charge is the same on each sphere—both exert a force of equal magnitude on a third charged object. In other words, each sphere gets half the original charge. By sharing charge among several identical metal spheres, Coulomb could produce spheres whose charge was one-half, one-quarter, one-eighth, and so on of the original charge (**Figure 22.26**).

By thus varying the charges on spheres A and B of his apparatus, Coulomb found that the **electric force** is proportional to the charge on each sphere. We can summarize these findings in one equation, called **Coulomb's law,** which gives the magnitude of the electric force exerted by two charged particles separated by a distance r_{12} and carrying charges q_1 and q_2:

$$F_{12}^E = k\frac{|q_1|\ |q_2|}{r_{12}^2}. \tag{22.1}$$

As we shall see in Chapter 27, the interaction between charged particles becomes more complicated when the particles are not at rest. For this reason the force in Coulomb's law is sometimes called the *electrostatic* force and the branch of physics that deals with stationary distributions of charge is called *electrostatics*.

If the positions of the two charged particles are given by the vectors $\vec{r}_1$ and $\vec{r}_2$, respectively, then the distance between them is $r_{12} = |\vec{r}_2 - \vec{r}_1|$. The value of the constant of proportionality k depends on the units used for charge, force, and length. The absolute-value signs around the charges in Eq. 22.1 are necessary because q_1 and q_2 can be negative but the *magnitude* F_{12}^E of the electric force must always be positive.

Coulomb's law bears a striking resemblance to Newton's law of gravity (Eq. 13.1):

$$F_{12}^G = G\frac{m_1 m_2}{r_{12}^2}. \tag{22.2}$$

Why these two laws have the same mathematical form remains a mystery. The main differences between the two are that mass is always positive but electrical charge can be positive or negative, which means that the gravitational force is always attractive but the electric force can be attractive or repulsive.

The derived SI unit of charge, called the **coulomb** (C), is the quantity of electrical charge transported in 1 s by a current of 1 ampere, a quantity and unit we shall define in Chapter 27. One coulomb is equal to the magnitude of the charge on about 6.24×10^{18} electrons. Conversely, the magnitude of the charge of the electron is

$$e = 1/6.24 \times 10^{18}\ \text{C} = 1.60 \times 10^{-19}\ \text{C}. \tag{22.3}$$

The charge on any object comes in only whole-number multiples of this elementary charge:

$$q = ne \quad (n = 0,\ \pm 1,\ \pm 2,\ \pm 3, \ldots). \tag{22.4}$$

This means that an object can have charge $q = 0$, $q = +7e$, $q = -4e$, and so forth, but not, for instance, $+1.2e$. Because the elementary charge is very small, the fact that charge exists only as whole-number multiples of the elementary charge isn't noticeable under ordinary circumstances. For example, running a comb through your hair easily gives the comb a surplus of about 10^{12} electrons, and the quantity of electrons flowing through

Figure 22.25 Schematic diagram of Coulomb's apparatus for measuring the electric force between two charged spheres.

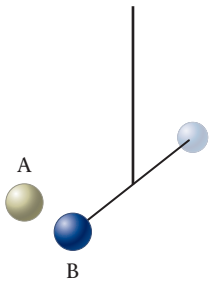
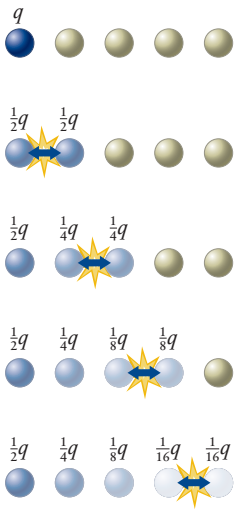


Figure 22.26 By successively allowing a charged sphere to touch an initially uncharged neighbor, we can distribute an amount of charge in ever-lessening amounts over a number of spheres.



a 100-W light bulb each second is about 10^{19}. These are such large numbers that the fact that charge comes in only whole-number multiples of the elementary charge normally remains unnoticed.

Using the coulomb as the unit of charge, we can determine the value of k in Eq. 22.1 experimentally by measuring the force between two known charged particles separated by a known distance:

$$k = 9.0 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2. \tag{22.5}$$

The value of this constant shows how large a unit the coulomb is: Two particles, each carrying a charge of 1 C, separated by 1 m exert on each other a force of 9 billion newtons—equal to the gravitational force exerted by Earth on several dozen loaded supertankers! It is very difficult to build up a charge of this magnitude on all but very large objects because things get ripped apart by the enormous forces. The largest accumulations of charge we know of occur in the atmosphere: Large clouds that accumulate a charge of about 50 C discharge through the air to Earth, causing lightning.

EXAMPLE 22.2 Gravity versus electricity

Compare the magnitudes of the gravitational and electric forces exerted by the nucleus of a hydrogen atom—a single proton ($m_p = 1.7 \times 10^{-27}$ kg)—on an electron ($m_e = 9.1 \times 10^{-31}$ kg) when the two are 0.50×10^{-10} m apart.

❶ **GETTING STARTED** For simplicity, I assume I can treat the proton and electron as particles. I also assume they are at rest so that I can use the principles of electrostatics.

❷ **DEVISE PLAN** I can use Eq. 22.2 to calculate the magnitude of the gravitational force and Eq. 22.1 to calculate the magnitude of the electric force.

❸ **EXECUTE PLAN**

$$F_{pe}^G = G\frac{m_p m_e}{r_{pe}^2}$$

$$= (6.7 \times 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2)\frac{(9.1 \times 10^{-31} \text{ kg})(1.7 \times 10^{-27} \text{ kg})}{(0.50 \times 10^{-10} \text{ m})^2}$$

$$= 4.1 \times 10^{-47} \text{ N}$$

and

$$F_{pe}^E = k\frac{|q_p|\,|q_e|}{r_{pe}^2}$$

$$= (9.0 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2)\frac{(1.6 \times 10^{-19} \text{ C})(1.6 \times 10^{-19} \text{ C})}{(0.50 \times 10^{-10} \text{ m})^2}$$

$$= 9.2 \times 10^{-8} \text{ N}.$$

The electric force exerted by the proton on the electron is $(9.2 \times 10^{-8} \text{ N})/(4.1 \times 10^{-47} \text{ N}) \approx 10^{39}$ times greater than the gravitational force exerted by the proton on the electron. ✔

❹ **EVALUATE RESULT** Table 7.1 gave the electromagnetic interaction for two protons as 36 orders of magnitude stronger than the gravitational interaction. Here we considered a proton and an electron—and found the electric force to be 39 orders of magnitude greater. An electron has about 3 orders of magnitude less mass than a proton, so it makes sense that the gravitational force on it is 3 orders of magnitude less.

EXAMPLE 22.3 Comb electricity

(*a*) A 0.020-kg plastic comb acquires a charge of about -1.0×10^{-8} C when passed through your hair. What is the magnitude of the electric force between two such combs held 1.0 m apart after being passed through your hair? (*b*) If two identical 0.020-kg combs carry one surplus electron for every 10^{11} electrons in the combs, what is the magnitude of the electric force between these combs held 1.0 m apart?

❶ **GETTING STARTED** Both parts of the problem require me to calculate the magnitude of the electric force between the combs. If I treat the combs as particles, I can use Eq. 22.1 to calculate this force.

❷ **DEVISE PLAN** To calculate the magnitude of the electric force between two charged objects, I need to know the charge on each object and their separation distance. I know these data for part *a*: $q_1 = q_2 = -1.0 \times 10^{-8}$ C and $r_{12} = 1.0$ m, where the subscripts 1 and 2 denote the two combs. For part *b* I am given only the separation distance, and so I need to determine the charge on each comb.

I am given the fraction of electrons added, and I know the charge on one electron. So, to determine the charge on each comb, I need to determine how many electrons each comb contains. The number of electrons in each comb is equal to the number of protons in the comb: $N_e = N_p$. I am given the mass of the comb and I know that the mass is determined by the protons and neutrons in all the atoms making up the comb (the electrons contribute very little). Given that the protons and neutrons have almost identical mass ($m_p = m_n = 1.7 \times 10^{-27}$ kg), I can determine the number N of protons and neutrons by dividing the mass of

QUANTITATIVE TOOLS

the comb by m_p. Given that most atoms contain roughly equal numbers of protons and neutrons, I can say that the number of protons is $N_p \approx N/2$.

3 EXECUTE PLAN (*a*) Substituting the values given into Eq. 22.1, I get

$$F_{12}^E = k\frac{|q_1|\ |q_2|}{r_{12}^2}$$
$$= (9.0\times10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)\frac{(1.0\times10^{-8}\ \text{C})(1.0\times10^{-8}\ \text{C})}{(1.0\ \text{m})^2}$$
$$= 9.0\times10^{-7}\ \text{N}. \checkmark$$

(*b*) The number of protons plus neutrons in the comb is

$$N = \frac{0.020\ \text{kg}}{1.7\times10^{-27}\ \text{kg}} = 1.2\times10^{25},$$

and so $N_p \approx N/2 = 6\times10^{24}$. The number of electrons is equal to the number of protons, and so there are 6×10^{24} electrons in each comb to begin with. Adding one surplus electron for every 10^{11} electrons means adding $(6\times10^{24})/(1\times10^{11}) = 6\times10^{13}$ electrons to each comb; these electrons carry a combined charge of $(6\times10^{13})(-1.6\times10^{-19}\ \text{C}) = -9.6\times10^{-6}\ \text{C}$. The magnitude of the repulsive electric force between the combs is then

$$F_{12}^E = k\frac{|q_1|\ |q_2|}{r_{12}^2}$$
$$= (9.0\times10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)\frac{(9.6\times10^{-6}\ \text{C})(9.6\times10^{-6}\ \text{C})}{(1.0\ \text{m})^2}$$
$$\approx 1\ \text{N}. \checkmark$$

4 EVALUATE RESULT My answer to part *a* is a force too small to be felt, which is what I expect based on experience (two combs passed through hair don't exert an appreciable force on each other). In contrast, my answer to part *b* is phenomenally large for an electric force. The magnitude of the initial acceleration acquired by the combs would be $a_1 = F_{21}^E/m_1 = (1\ \text{N})/(0.020\ \text{kg}) = 50\ \text{m/s}^2$, or about five times the acceleration due to gravity! Even though the fraction of electrons removed—one in 100 billion—is very small, the factor k in Eq. 22.1 is so great that the resulting force is also great. Indeed, I learned in Table 7.1 that the electromagnetic interaction is 36 orders of magnitude stronger than the gravitational interaction, so my answer is not unreasonable.

22.16 Two identical conducting spheres, one carrying charge $+q$ and the other carrying charge $+3q$, are initially held a distance d apart. The spheres are allowed to touch briefly and then returned to separation distance d. Is the magnitude of the force they exert on each other after the touching greater than, smaller than, or the same as the magnitude of the force they exerted on each other before the touching?

Like the gravitational force, the electric force is *central;* that is, its line of action is along the line connecting the two interacting charged particles. Consider, for example, the two particles carrying charges q_1 and q_2 shown in **Figure 22.27a**. The vector $\vec{r}_{12} \equiv \vec{r}_2 - \vec{r}_1$ gives the position of particle 2 relative to particle 1; this vector points from particle 1 to particle 2. We can define a unit vector pointing in this direction by dividing the vector $\vec{r}_{12}$ by its magnitude:

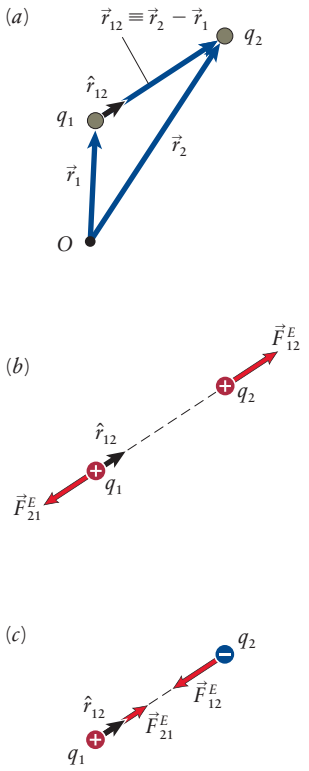
$$\hat{r}_{12} \equiv \frac{\vec{r}_2 - \vec{r}_1}{r_{12}}. \tag{22.6}$$

Depending on the algebraic sign of the charges, the electric force can be attractive or repulsive. For like charges ($q_1q_2 > 0$), the force is repulsive. In this case, the force $\vec{F}_{12}^E$ exerted by particle 1 on particle 2 points in the same direction as the unit vector $\hat{r}_{12}$ (Figure 22.27*b*). For opposite charges ($q_1q_2 < 0$), the force is attractive, and so $\vec{F}_{12}^E$ points in the direction opposite the direction of $\hat{r}_{12}$ (Figure 22.27*c*). In either case, $\vec{F}_{12}^E$ can be written in the form

$$\vec{F}_{12}^E = k\frac{q_1q_2}{r_{12}^2}\hat{r}_{12}. \tag{22.7}$$

Because $r_{12} = r_{21}$ and because $\hat{r}_{21}$ points in the direction opposite the direction of $\hat{r}_{12}$, the force $\vec{F}_{21}^E$ exerted *by* particle 2 *on* particle 1, which

Figure 22.27 (*a*) Position vectors for two charged particles. (*b*) Repulsive forces exerted on each other by two particles carrying like charges. (*c*) Attractive forces exerted on each other by two particles carrying opposite charges.



points in the direction opposite the direction of $\vec{F}_{12}^E$, is obtained by simply switching the indices 1 and 2 in Eq. 22.7:

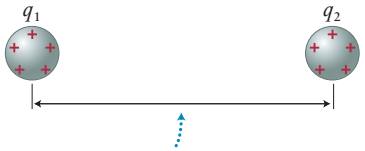
$$\vec{F}_{21}^E = k\frac{q_2q_1}{r_{21}^2}\hat{r}_{21} = k\frac{q_1q_2}{r_{12}^2}(-\hat{r}_{12}) = -\vec{F}_{12}^E, \tag{22.8}$$

as we would expect for an interaction pair (see Eq. 8.15). Equation 22.7 is the vectorial form of Eq. 22.1.

22.17 Using your knowledge about work and potential energy, determine whether the potential energy of a closed system of two charged particles carrying like charge increases, decreases, or stays the same when the distance between the two is increased. Repeat for two particles carrying opposite charge.

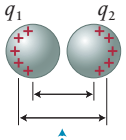
Figure 22.28 The reason Coulomb's law does not apply in a strict sense to macroscopic charged objects. The law is approximately correct if the objects are far apart relative to their radii.

(*a*) Charged spheres separated by a distance large compared to the sphere radii



Distance between centers of charge distributions same as distance between sphere centers

(*b*) Charged spheres separated by a distance small compared to the sphere radii



Charge repulsion causes distance between centers of charge distributions to differ from distance between sphere centers.

Before going on, I should mention a limitation to Coulomb's law. Strictly speaking, it is applicable only to charged particles. This is so because the distance r_{12} is well defined only when the size of the charged objects is negligibly small compared with their separation distance. When the charged objects are not particles, the distance r_{12} is not equal to the center-to-center distance. You can see why with the help of Figure 22.28. In **Figure 22.28*a***, the charge is distributed uniformly over the surface of each of two widely separated metal spheres. The way in which a collection of charge carriers is spread out over a macroscopic object is called a **charge distribution.** Because the charge distributions over the metal spheres in Figure 22.28*a* are uniform, the center of each charge distribution coincides with the center of the sphere, and so r_{12} is well defined. When we bring the spheres close together, as in Figure 22.28*b*, the like charge carriers repel one another and move to the far side of each sphere. Now the centers of the charge distributions no longer coincide with the spheres' centers, so r_{12} (the center-to-center distance of the two charge distributions) is not simply the distance separating the centers of the two conductors.

22.18 (*a*) Is the magnitude of the electric force between the two conducting spheres in Figure 22.28*b* greater or smaller than that obtained from Coulomb's law, which assumes the charge is concentrated at the center of each sphere? (*b*) Is the answer to part *a* the same if the charge on one of the conductors is negative instead of positive?

22.6 Forces exerted by distributions of charge carriers

Coulomb's law deals only with *pairs* of charged objects. To calculate the force exerted by an assembly of objects carrying charges q_2, q_3, q_4, . . . on an object 1 carrying a charge q_1, we take the vector sum of all the forces exerted on object 1 by each of the other charged objects independently:

$$\Sigma\vec{F}_1^E = \vec{F}_{21}^E + \vec{F}_{31}^E + \vec{F}_{41}^E + \cdots, \tag{22.9}$$

where each term is given by Coulomb's law:

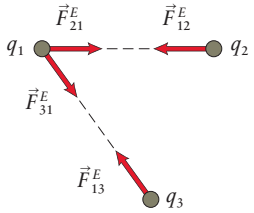
$$\Sigma\vec{F}_1^E = k\frac{q_2q_1}{r_{21}^2}\hat{r}_{21} + k\frac{q_3q_1}{r_{31}^2}\hat{r}_{31} + k\frac{q_4q_1}{r_{41}^2}\hat{r}_{41} + \cdots. \tag{22.10}$$

In other words, we calculate the force exerted by object 2 on object 1, then calculate the force exerted by object 3 on object 1, and so forth, and then add the forces. This means that if we know the details of some distribution of charged objects, we can calculate the force exerted by this distribution of charged objects on a single charged particle. For distributions that contain large numbers of charged objects, the summation can be

accomplished via an integration. We will limit our discussion here to simple cases involving only a few charged objects.

22.19 **Figure 22.29** shows how a charged particle 1 interacts with two other charged particles 2 and 3. Determine the direction of the vector sum of the electric forces exerted on particle 2.

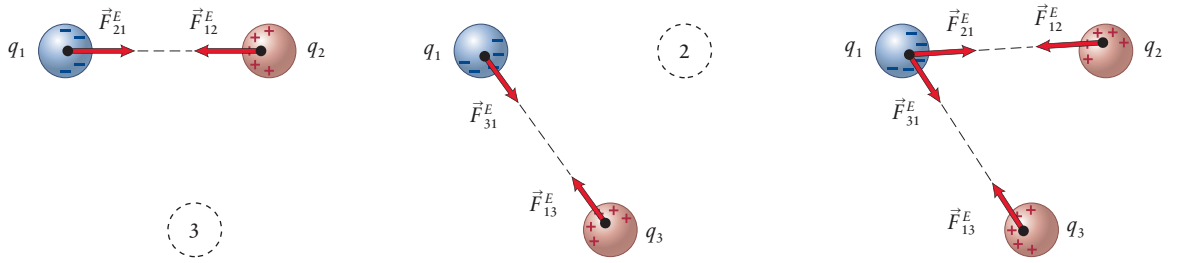
Figure 22.29 Forces exerted by two charged particles 2 and 3 on charged particle 1.



The basic limitation of Coulomb's law continues to apply when we are analyzing collections of charged objects: Eq. 22.10 is valid only for charged *particles*, not for charged extended bodies. Suppose, for example, that we replace each charged particle in Figure 22.29 by a conducting sphere carrying the same charge as each particle. Let us first consider just the interaction between spheres 1 and 2. When these oppositely charged spheres are placed near each other, the charge carriers on the two spheres rearrange themselves to be as close as possible to each other (**Figure 22.30***a*). A similar type of rearrangement takes place when just spheres 1 and 3 are placed near each other (Figure 22.30*b*). When all three spheres are placed near one another, the positive charge on sphere 3 pushes the positive charge on sphere 2 up and pulls the negative charge on sphere 1 down (Figure 22.30*c*). Consequently, the forces that the spheres exert on one another are not the same as the forces exerted by the individual pairs (compare the forces in Figures 22.29 and 22.30).

Figure 22.30

(*a*) Sphere 1 interacts with just sphere 2 (*b*) Sphere 1 interacts with just sphere 3 (*c*) Sphere 1 interacts with both spheres



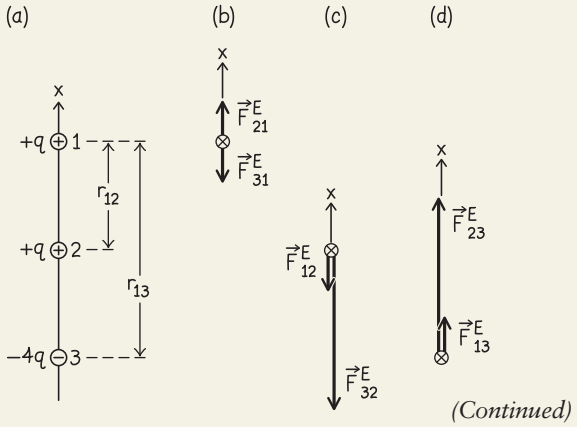
EXAMPLE 22.4 Electric tug of war

You are given three charged particles. Particles 1 and 2 carry charge $+q$ and particle 3 carries charge $-4q$. (*a*) Determine the relative values of the separation distances r_{12} and r_{13} when the three particles are arranged along a straight line in such a way that the vector sum of the forces exerted on particle 1 is zero. (*b*) With the particles arranged this way, are the vector sums of the forces exerted on particles 2 and 3 also zero?

❶ **GETTING STARTED** Particle 2 exerts a repulsive force on particle 1; particle 3 exerts an attractive force on particle 1. In order for the two forces exerted on particle 1 to cancel, particles 2 and 3 must be on the same side of particle 1. Because the magnitude of the charge on particle 2 is smaller than that of the charge on particle 3, the distance r_{12} must be shorter than the distance r_{13}.

❷ **DEVISE PLAN** I choose my x axis vertically up along the line defined by the particles (**Figure 22.31***a*). To determine the relative values of the separation distances r_{12} and r_{13} when the vector sum of the forces exerted on particle 1 is zero, I draw a free-body diagram for particle 1 (Figure 22.31*b*). The magnitude of each force exerted on this particle is given by

Figure 22.31



(Continued)

QUANTITATIVE TOOLS

Eq. 22.1, and so by setting the sum of the x components of the forces equal to zero, I have an expression containing r_{12} and r_{13} and I can manipulate the expression to get the relative values of r_{12} and r_{13}.

❸ **EXECUTE PLAN** (*a*) The x components of the two forces in Figure 22.31*b* must add to zero, so

$$\begin{aligned}\Sigma F_{1x} &= F_{21x}^E + F_{31x}^E \\ &= +k\frac{|q_1||q_2|}{r_{12}^2} - k\frac{|q_1||q_3|}{r_{13}^2} = k\frac{qq}{r_{12}^2} - k\frac{(q)(4q)}{r_{13}^2} \\ &= +k\frac{q^2}{r_{12}^2} - k\frac{4q^2}{r_{13}^2} = 0.\end{aligned}$$

I therefore must have $\dfrac{q^2}{r_{12}^2} = \dfrac{4q^2}{r_{13}^2}$,

or $r_{13}^2 = 4r_{12}^2$, and so $r_{13} = 2r_{12}$. ✔

(*b*) The forces exerted by particles 1 and 3 on particle 2 both point in the negative x direction (Figure 22.31*c*) and so cannot sum to zero. The forces exerted by particles 1 and 2 on particle 3 both point in the positive x direction (Figure 22.31*d*) and so cannot sum to zero. ✔

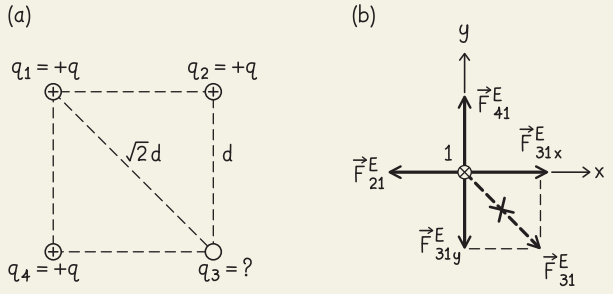
❹ **EVALUATE RESULT** My answer for part *a* makes sense because the force exerted by each particle is inversely proportional to the square of the separation distance, and this force varies directly with the quantity of charge. Because the charge on particle 3 is four times greater than that on particle 2, the square of the separation distance r_{13} must be four times r_{12}, and so $r_{13} = 2r_{12}$, as I found.

EXAMPLE 22.5 Electrostatic equilibrium

Consider four charged particles placed at the corners of a square whose sides have length d. Particles 1, 2, and 4 carry identical positive charges. In order for the vector sum of the forces exerted on particle 1 to be zero, what charge must be given to particle 3, which is in the corner diametrically opposite particle 1?

❶ **GETTING STARTED** I begin by making a sketch of the situation (**Figure 22.32*a***). Because particles 1, 2, and 4 all carry identical positive charge, I write $q_1 = q_2 = q_4 = +q$. The separation between neighboring particles on the square is d. The separation between particles 1 and 3 is $\sqrt{2}d$.

Figure 22.32



❷ **DEVISE PLAN** To determine the charge needed on particle 3, I must determine the electric force magnitude F_{31}^E needed to yield a zero vector sum of forces exerted on particle 1. I should therefore draw a free-body diagram for particle 1 and work out the vector sum.

❸ **EXECUTE PLAN** In my free-body diagram (Figure 22.32*b*), I choose my y axis pointing up and x axis pointing to the right. The force $\vec{F}_{21}^E$ is repulsive and so points in the negative x direction; the force $\vec{F}_{41}^E$ is repulsive and points in the positive y direction. To make the vector sum of the forces exerted on 1 zero, $\vec{F}_{31}^E$ must be such that its components cancel $\vec{F}_{21}^E$ and $\vec{F}_{41}^E$. Along the x axis, I therefore have

$$\Sigma F_{1x} = F_{21x}^E + F_{31x}^E + F_{41x}^E = -F_{21}^E + F_{31}^E \cos 45° + 0 = 0,$$

so $F_{21}^E = F_{31}^E \cos 45°$. Substituting the Coulomb's law expressions for these two force magnitudes, I have

$$k\frac{|q_2||q_1|}{d^2} = k\frac{|q_3||q_1|}{(d\sqrt{2})^2}\cos 45°$$

$$\frac{q^2}{d^2} = \frac{|q_3|q}{2\sqrt{2}d^2}$$

$$|q_3| = 2\sqrt{2}\,q.$$

From Figure 22.32*b*, I also know that $\vec{F}_{31}^E$ must be an attractive force, so

$$q_3 = -2\sqrt{2}q. \checkmark$$

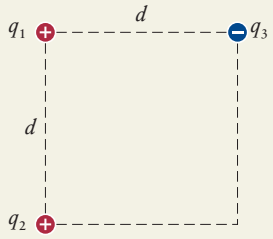
❹ **EVALUATE RESULT** My result indicates that the charge on particle 3 is greater than q. This makes sense because the attractive force exerted by this particle must balance the forces exerted by particles 2 and 4, which are both closer to particle 1. To obtain my answer I solved only for the x component of $\vec{F}_{31}^E$ and did not consider the y component. However, my free-body diagram shows that the magnitudes of the components along the y axis are the same as those along the x axis, so analyzing the y components would have given me the same result.

QUANTITATIVE TOOLS

EXAMPLE 22.6 Electric trajectory

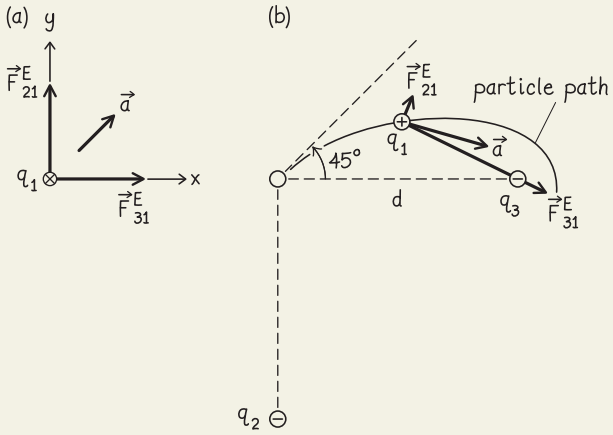
Consider the arrangement of charged particles shown in Figure 22.33. The charge magnitudes are the same in all three cases, but q_1 and q_2 are positive and q_3 is negative. Sketch the trajectory of particle 1 if it is released while particles 2 and 3 are held fixed. Ignore any gravitational force exerted by Earth on the particle.

Figure 22.33 Example 22.6



❶ **GETTING STARTED** I begin by drawing a free-body diagram for particle 1, choosing the x axis to the right and the y axis up (Figure 22.34*a*). Because the charges on the two particles have the same magnitude and because the separation distances are the same, the magnitudes F_{21}^E and F_{31}^E are the same.

Figure 22.34



❷ **DEVISE PLAN** The direction in which particle 1 accelerates is determined by the vector sum of the forces exerted on it, and I can determine this direction from my free-body diagram. As the particle moves along its trajectory, however, the forces exerted on it change in both direction and magnitude, and so I must consider these changes in formulating my answer.

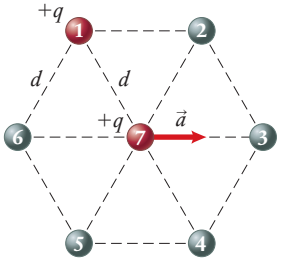
❸ **EXECUTE PLAN** Because the magnitudes F_{21}^E and F_{31}^E are the same, the vector sum of the two forces exerted on particle 1 bisects the angle between the two forces, and so the initial acceleration of 1 (which points in the same direction as the vector sum of forces) points up and to the right at an angle of 45°, as illustrated in Figure 22.34*a*. As particle 1 moves in the direction indicated in Figure 22.34*a*, $\vec{F}_{21}^E$ and $\vec{F}_{31}^E$ change in both direction and magnitude. The magnitude of $\vec{F}_{21}^E$ decreases because the distance between 1 and 2 increases, and the magnitude of $\vec{F}_{31}^E$ increases because the distance between 1 and 3 decreases. The direction of the acceleration of particle 1 is the same as the direction of the vector sum of these two forces. The resultant path of motion is qualitatively illustrated in Figure 22.34*b*. ✔

❹ **EVALUATE RESULT** My sketch makes sense: Particle 1 first moves up and to the right because it is repelled by particle 2 and attracted by particle 3. As it moves away from 2 and approaches 3, the effect of the attraction increases and so the trajectory curves and heads toward 3.



22.20 Seven small metal spheres are arranged in a hexagonal pattern as illustrated in Figure 22.35. Spheres 1 and 7 carry equal amounts of positive charge; the other spheres are uncharged. (*a*) To give sphere 7 an acceleration $\vec{a}$ that points to the right, what (single) other sphere must be charged? There may be more than one possibility. (*b*) What are the sign and magnitude of that charge?

Figure 22.35 Checkpoint 22.20.



Chapter Summary

Electric interactions (Sections 22.1, 22.2, 22.5)

Concepts An **electric interaction** is a long-range interaction between objects that carry a surplus **electrical charge.** The electric force is sometimes called the *electrostatic* force because interactions between charged objects become more complicated for objects that are not at rest (see Chapter 27).

There are only two types of charge: **positive charge** and **negative charge.** Objects that carry like charges repel each other; objects that carry opposite charges attract each other. Any microscopic object that carries electrical charge, such as an electron or ion, is called a **charge carrier.**

The SI unit of charge is the **coulomb** (C), which is the magnitude of the charge carried by about 6.24×10^{18} electrons. It is also equal to the amount of charge transported in 1 s by a steady current of 1 ampere (see Chapter 27).

Quantitative Tools **Coulomb's law** states that if particle 1 carrying charge q_1 and particle 2 carrying charge q_2 are a distance r_{12} apart, the **electric force** $\vec{F}_{12}^{E}$ exerted by particle 1 on particle 2 is

$$\vec{F}_{12}^{E} = k\frac{q_1 q_2}{r_{12}^2}\hat{r}_{12}, \tag{22.7}$$

where the unit vector $\hat{r}_{12}$ points from particle 1 to particle 2 and the constant k is

$$k = 9.0 \times 10^9\ \mathrm{N \cdot m^2/C^2}. \tag{22.5}$$

The **elementary charge** e is the smallest observed amount of charge and is equal to the magnitude of the charge carried by one electron:

$$e = 1.60 \times 10^{-19}\ \mathrm{C}. \tag{22.3}$$

The charge q carried by any object comes in whole-number multiples of e:

$$q = ne \quad (n = 0, \pm 1, \pm 2, \pm 3, \ldots). \tag{22.4}$$

Movement of charge carriers (Sections 22.3, 22.4)

Concepts Charge carriers can flow through an **electrical conductor** but cannot readily flow through an **electrical insulator.** Charge carriers can be transferred from one nonconductor to another only by bringing the nonconductors into contact with each other, and the carriers remain near the spot at which they were deposited.

We **ground** an object by connecting it electrically to Earth. Grounding permits the exchange of charge carriers with Earth, a huge reservoir of charge carriers. A charged, conducting object that is grounded will retain no surplus of either type of charge, assuming no other nearby electrical influences.

An **ion** is an atom or molecule that has lost or gained one or more electrons.

According to the principle of **conservation of charge,** electrical charge can be created or destroyed only in identical positive-negative pairs, with the result that the charge of a closed system always remains constant.

Polarization, which is the separation of charge carriers in an object, allows neutral objects to interact electrically.

Charging by induction is a method of using a charged object to charge a neutral object without the two objects touching each other.

Forces due to multiple charges (Section 22.6)

Quantitative Tools The electric force exerted on a charged particle 1 by any number of other charged particles 2, 3, . . . is

$$\Sigma\vec{F}_1^E = k\frac{q_2q_1}{r_{21}^2}\hat{r}_{21} + k\frac{q_3q_1}{r_{31}^2}\hat{r}_{31} + \cdots. \qquad (22.10)$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

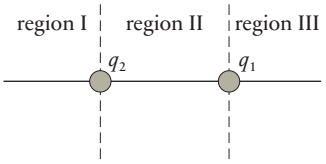
22.1 Static electricity

1. A piece of tape is pulled from a spool and lowered toward a 100-mg scrap of paper. Only when the tape comes within 8.0 mm is the electric force magnitude great enough to overcome the gravitational force exerted by Earth on the scrap and lift it. Determine the magnitude and direction of the electric force exerted by the tape on the paper at this distance. •
2. A baseball bat is placed on a table such that the narrow end of the bat hangs over the edge. Two freshly pulled strips of tape are hung from opposite sides of the overhanging end of the bat. The strips repel each other so that they hang in the air in an inverted V shape, with an angle of about 50° between the strips. Describe at least three things you could do to reduce the angle between the hanging strips. •
3. A plastic spring with spring constant $k = 450$ N/m has a relaxed length of 0.100 m. The spring is positioned vertically on a table, and a charged plastic 1.00-kg sphere is placed on the top end of the spring. Another charged object is suspended above the sphere without making contact. If the length of the spring is now 0.0950 m, what are the magnitude and direction of the electric force exerted on the sphere? ••

22.2 Electrical charge

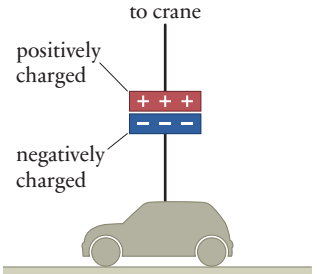
4. Two small objects carry the same amount of charge, and the magnitude of the electric force exerted by one object on the other is 0.10 N when they are held 30 mm apart. If the force magnitude increases when the objects are released and free to move, are the objects of the same charge or opposite charge? •
5. How might the results from electrostatic experiments with tape strips described in Section 22.2 differ from those actually observed if there were three types of electrical charges in the universe? ••
6. If you must do positive work to bring a charged balloon toward a negatively charged sphere, is the charge on the balloon positive or negative? ••
7. Suppose you have three identical, electrically neutral pieces of tape, A, B, and C. (*a*) Describe how you could make two of them repel each other. (*b*) Is it possible for all three to repel one another? ••
8. Suppose that in Figure P22.8, $|q_2| > |q_1| > 0$ but these particles may carry charges of the same or opposite type. A third particle carrying charge q_3 is to be placed in this system along the axis such that the vector sum of electric forces on the third particle is zero. In which region—I, II, or III—might this be possible if (*a*) q_1 and q_2 have the same sign and (*b*) q_1 and q_2 have opposite signs? (*c*) Under what circumstances can the third particle be placed in region I and have the vector sum of electric forces exerted on it be zero? ••

Figure P22.8



9. A salvage yard operator wants to make an electric force linkage for his wrecker crane as a safety device, so that when a vehicle being lifted is too massive, it will break the linkage before the force exerted by the vehicle topples the crane. His plan is to attach two metal plates to the crane cable as shown in Figure P22.9 and connect the plates to the crane's power system. There is to be a positive charge on one plate and a negative charge on the other, so that the two stick together because of the attractive electric force. Discharge is prevented by spraying a layer of plastic on each plate. He needs to know what diameter to make the plates so that they come apart when the vehicle being lifted has a mass greater than 2000 kg. On the micro level, the electric force is between surface atoms carrying positive charge on one plate and surface atoms carrying negative charge on the other plate. The electric force for each atom pair has a magnitude of 2.3×10^{-12} N, and each atom occupies an area of 3×10^{-21} m^2. (*a*) Before doing any calculations, estimate the necessary plate surface area. Which is this surface area comparable to: that of a pinhead, a fingernail, a music compact disc, a dinner plate, or a manhole cover? (*b*) Calculate the necessary area. How accurate was your estimate as to which object the plate area is comparable to? •••

Figure P22.9



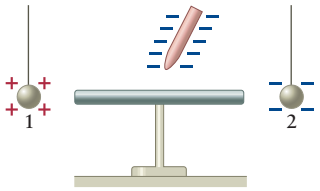
22.3 Mobility of charge carriers

10. Which of the following phenomena are due to the electric interaction: (*a*) dissolution of salt in water, (*b*) surface tension in water, (*c*) the elliptical orbit of comets, (*d*) the binding of protons in an atomic nucleus, (*e*) "traction" between tires and pavement. •
11. In an airplane, the body of the aircraft gets charged if friction occurs between it and the air, which contains dust particles. As there is no possibility of grounding through a copper wire on the ground, the aircraft contains sharp wicks. Explain their role in protecting the aircraft from damage due to accumulation of excessive charge. •
12. You have two rods A and B, both made of a nonconducting material. Initially, neither rod is charged. A sphere carrying 100 units of surplus positive charge is placed close to the left end of rod A. One end of rod B is touched to the right end of rod A for some time interval and then removed. What is the final charge on rod B? •
13. Some survivors of lightning strikes have reported feeling their hair standing on end shortly before the strike. What causes the hair to behave in this way? •
14. A closed system consists of a neutron and a positron. The neutron eventually decays. After decay, what are the most likely particles in the system, and what is the charge in the system? (Use the principle of conservation of charge.) ••
15. Two identical 500-mg paper clips, 1 and 2, dangle from two strings, the top ends of which are tied to the same nail. The clips are charged in such a way that the two strings hang with an angle of 32.0° between them. (*a*) Draw a free-body diagram for clip 1, which is to the left of clip 2. (*b*) What are the magnitude and direction of the electric force exerted by one clip on the other? (*c*) If the two clips were brought in contact with each other for an instant and then released, might they discharge into each other and no longer repel? ••
16. A metal sphere is grounded and an electrically charged rod is held close to it. The sphere acquires a charge q. The grounding wire is removed, and the sphere is brought in contact with another uncharged identical metal sphere. (*a*) What is the charge on each sphere after they are separated? (*b*) One of the two spheres is grounded again, and the procedure is repeated. What is the charge on each sphere now? (*c*) What is the maximum charge that can be put on each sphere? ••
17. At separation distance r_1, charged Ping-Pong balls A and B repel each other with an electric force of magnitude 0.40 N. At separation distance r_2, charged Ping-Pong balls A and C repel each other with a force of magnitude 1.4 N. Maintaining the separation distances r_1 and r_2, how should you arrange B and C around A (*a*) to create the maximum electric force exerted on A and (*b*) to create the minimum electric force exerted on A? ••
18. A metal sphere is grounded via a single wire. One end of a plastic rod is given a positive charge and held near the sphere. (*a*) If the ground wire is removed while the rod and sphere are near each other but have not touched, is there an electric force between the two? Why or why not? (*b*) If the rod end is brought in contact with the sphere and then the ground wire is removed, is there an electric force between rod and sphere after they are separated? Why or why not? •••

22.4 Charge polarization

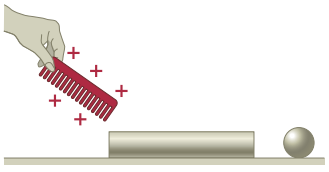
19. If you rub a balloon with hair and bring it close to a stream of water in a wash basin, the water gets attracted toward the balloon. Can you predict the charge on the stream of water? •
20. Three small pieces of electrically neutral paper are placed far apart from one another on a table. One piece is given a negative electric charge, one is given a positive charge, and one is left neutral. They all look identical, and in order to tell which is which, a student runs a comb through her hair and waves the comb over each piece. She observes that one is strongly attracted to the comb; she labels this A. She observes that another is weakly attracted to the comb; she labels this B. The last piece, which is repelled by the comb; she labels C. Which piece is positive, which is negative, and which is neutral? •
21. The electrically neutral metal rod in Figure P22.21 stands between two charged pith balls, and then a negatively charged rubber rod is scraped along the center of the metal rod. Draw the free-body diagram for each ball after the rubber rod is taken away. •

Figure P22.21



22. A charged sphere with 1×10^8 units of negative charge is brought near a neutral metal rod. The half of the rod closer to the sphere has a surplus charge of 1×10^4 units. What is the magnitude of the charge on the half of the rod farther from the sphere? Is this charge positive or negative? •
23. Near one end of an electrically neutral aluminum rod lying on a table made of nonconducting material, you hold a comb that is also made of a nonconducting material but carries a positive charge (Figure P22.23). Does a small steel sphere located near the other end of the rod roll toward the rod, roll away from the rod, or stay where it is if the sphere is (*a*) electrically neutral and (*b*) positively charged? ••

Figure P22.23



24. Two metal rods A and B are initially lying end to end on a wooden table and are touching each other. A sphere carrying surplus positive charge is brought close to the middle of rod A. While the sphere is still in place, the rods are pulled apart by someone wearing gloves made of a fabric that is an electrical insulator. Then the sphere

is moved far away. (*a*) Do the two rods attract each other, do they repel each other, or is there no electric force between them? (*b*) Describe the surplus charge on each rod as positive, zero, or negative. ••

25. You have three pairs of spheres. Pair A consists of two conducting spheres, pair B of two nonconducting spheres, and pair C of one nonconducting sphere and one conducting sphere. If the same charge q is added to all spheres uniformly and in each pair the two spheres are separated by the same distance d, rate the magnitude of the electric force between each pair's spheres when q is (*a*) positive on both spheres, (*b*) negative on both, and (*c*) positive on one sphere in the pair and negative on the other. ••

26. Metal sphere A is hung from the ceiling by a long, thin string and given a positive charge. An identical sphere B is suspended nearby with an identical string and given a *negative* charge that has the same magnitude as the charge on A. The two spheres are 50 mm apart when in equilibrium. If an identical sphere C carrying no surplus charge is placed between them, does the distance between A and B increase, decrease, or remain the same? ••

27. You are an astrophysicist working on a problem involving Saturn's rings and trying to figure out why the smaller ice chunks in the rings stick to larger chunks. You hypothesize that as the chunks come out of the shadow of Saturn into the sunlight, the outside of the larger chunks heats up sooner than the center and this temperature difference creates a nonuniform distribution of the charged particles in the larger chunks. (Charge polarization due to temperature differences is known as the "thermoelectric effect.") Then, even though the chunk is electrically neutral, it has a charge separation. The small ice chunks presumably heat up uniformly all the way through and hence are not affected. Explain why small chunks gradually stick to larger ones. •••

22.5 Coulomb's law

28. By rubbing object A against B, if 10 electrons get transferred from A to B, calculate the electric force between the two objects if separated 10 cm apart. •

29. Two electrons are separated by 1.50 nm. What is the magnitude of the electric force each electron exerts on the other? Is this force attractive or repulsive? •

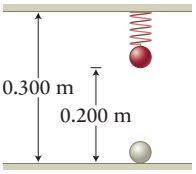
30. Two particles carrying charges q_1 and q_2 are separated by a distance r and exert an electric force of magnitude F^{E} on each other. If q_1 is doubled and q_2 is halved, what distance between them can keep the magnitude F^{E} constant? •

31. A particle carrying a -4.0-μC charge is located at the origin of an xy coordinate system. What are the components of the electric force exerted on a particle carrying a $+1.0$-μC charge if that particle is located (*a*) at $x = 10$ m, $y = 0$ and (*b*) at $x = 0$, $y = -6.0$ m? •

32. (*a*) By what distance do two objects carrying 1.0 C of charge each have to be separated before the electric force exerted on each object is 1.0 N? (*b*) How much charge should be placed on each of two small objects separated by 1.0 m so that the electric force exerted on each is 1.0 N? •

33. In a chamber 0.300 m high, sphere 1 carrying charge $q_1 = 1.00$ μC is suspended by a spring above a sphere 2 that carries charge q_2 and sits on the floor of the chamber (Figure P22.33). The relaxed length of the suspended spring is 70.0 mm, its spring constant is 30.0 N/m, and the mass of sphere 1 is 5.00 g. If sphere 1 comes to rest 0.200 m above the floor (and above sphere 2), what is the value of q_2? ••

Figure P22.33



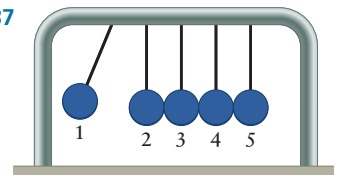
34. Two identical 9.60-g metal spheres (small enough to be treated as particles) are hung from separate 300-mm strings attached to the same nail in a ceiling. Surplus electrons are added to each sphere, and then the spheres are brought in contact with each other and released. Their equilibrium position is such that each string makes a 13.0° angle with the vertical. How many surplus electrons are on each sphere? ••

35. Suppose two particles carry a positive charge of 1.00 C for every kilogram of their mass. Compare the gravitational force and the electric force they exert on each other at a given separation. ••

36. Two identical metal spheres, one carrying a 12.0 μC charge and the other carrying a -32.0 μC charge, are initially 50 mm apart. They are brought in contact with each other and then moved back to their original positions. What is the numerical value of the ratio $F_{\mathrm{i}}^{E}/F_{\mathrm{f}}^{E}$, where F_{i}^{E} is the magnitude of the electric force between the spheres before they are brought in contact and F_{f}^{E} is the magnitude of that force after they are brought together and then moved back to their initial positions? ••

37. The device in Figure P22.37 is a so-called Newton's cradle. When you raise steel sphere 1 and then release it, its momentum is transferred to the other four steel spheres, and this momentum transfer causes sphere 5 to rise. When it falls, momentum is again transferred through the row of spheres, and sphere 1 rises and falls, and so on. Imagine modifying the device by positioning a charged wire in such a way that every time each end steel sphere rises, it touches the wire and acquires 10 units of charge. You lift sphere 1 and allow it to touch the wire for the first time and acquire 10 units of charge. You release the sphere and it hits the others, transferring its momentum and causing sphere 5 to rise. Sphere 5

Figure P22.37



touches the charged wire, descends, hits the row of spheres, and transfers its momentum so that sphere 1 rises again and touches the wire for the second time. If the net charge distributes uniformly over all balls, how many units of charge does sphere 1 carry (*a*) just before it touches the wire this second time and (*b*) just before it touches the wire after one more round of this momentum-transfer motion? ••

38. Two 1.00-kg spheres carry identical charges and are placed a distance d from each other. (*a*) How large should the charge on each sphere be so that the repulsive electric force between them balances the attractive gravitational force between them? (*b*) What is the smallest mass that can balance in this way, given that the smallest possible charge is $e = 1.602 \times 10^{-19}$ C? (*c*) Why doesn't the distance between the balls affect this calculation? ••

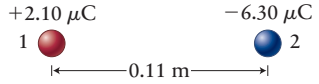
39. A charged oil drop that has a mass of 5.00×10^{-12} kg and carries 10 elementary charges is centered 50 mm below the center of an oppositely charged nonconducting sphere that has a radius of 5.0 mm. What uniformly distributed charge should the sphere be given so that the electric force it exerts on the drop causes the drop to float in the air for an instant? Treat the drop as a particle. ••

40. In the Bohr model, a hydrogen atom consists of an electron orbiting a one-proton nucleus. Furthermore, the angular momentum (see *Principles*, Chapter 11) of the electron is quantized, i.e., it equals $n\hbar$, where n is an integer and $\hbar$ a fundamental constant. Derive an algebraic expression for the relationship between the radius of the electron's orbit R_{orbit} and the fundamental constant $\hbar$. ••

41. You scrunch up three same-size sheets of aluminum foil to form three balls 1, 2, and 3, and stick each one onto the end of a wooden stake (which allows you to transport the balls without touching them). You place the stakes holding balls 1 and 2 an arm's length apart and then give each ball charge q. Call the electric force exerted by ball 1 on ball 2 $\vec{F}_{12}^{E}$. You then place charge $-2q$ on ball 3 and bring it first in contact with ball 1. Then you move it away from ball 1 and bring it in contact with ball 2. (*a*) Once you have moved ball 3 far away from balls 1 and 2, what is the magnitude, in terms of F_{12}^{E}, of the electric force between balls 1 and 2? (*b*) Has the direction of $\vec{F}_{12}^{E}$ changed? ••

42. (*a*) Draw and label the electric force exerted on each of the charged spheres in Figure P22.42 and compute their magnitudes. The centers of the spheres are separated by 0.11 m. (*b*) Repeat part *a*, drawing to the same scale, if the charge on the left sphere is $-6.30\ \mu$C. ••

Figure P22.42



43. Sphere A carries 6.0 nC of charge. It is placed 100 mm from sphere B, which carries 3.0 nC of charge. Assume the spheres are much smaller than their separation distance. (*a*) What is the magnitude of the electric force exerted on A? (*b*) Draw a figure showing the charges and the electric force vectors on each sphere. (*c*) Draw and label $\vec{r}_{\text{AB}}$ on your diagram. ••

44. You have two marbles, one carrying a uniformly distributed charge $q_{\text{p}} = +1.0\ \mu$C and the other carrying a uniformly distributed charge $q_{\text{n}} = -0.50\ \mu$C. (*a*) What is the magnitude of the electric force between them when they are held 100 mm apart? (*b*) What happens when the marbles are released and so are free to move? (*c*) If the marbles are metallic and share charge when they come in contact with each other, what is the magnitude of the electric force between them when they are again moved 100 mm apart, assuming no charge is transferred either to or from the surroundings of the two-marble system? ••

45. Two charged particles 3.0 m apart exert on each other an attractive electric force of magnitude 8.0×10^{-3} N. If the charge on the two particles sums to 6.0 μC, what is the charge on each particle? ••

46. A charged particle of 40-nC is located at the origin of the coordinate system. A second charged particle of 50 nC is located at (x, y) = (0, 2m) and another charged particle of 50 nC is located at (x, y) = (2m, 2m). Find the magnitude and direction of the electric force exerted by the first charged particle (*a*) on the second charged particle, and (*b*) on the third charged particle. ••

47. The electric force between two identical positively charged ions is 3.7×10^{-9} N when they are 0.50 nm apart. How many electrons are missing from each of the original atoms? ••

48. Two particles with identical positive charges separated by 2.5 cm are released from rest on a nonconducting, low-friction track. Immediately after release, one particle, which has a mass of 10 mg, has an acceleration of magnitude 5.0 m/s^2, while the other particle has an acceleration of magnitude 9.0 m/s^2. Find (*a*) the mass of the second particle and (*b*) the charge on each particle. ••

49. Two identical lightweight conducting balls are suspended by thin strings of identical length from a common point. One ball is given a charge q, and both are constrained in their motion by only the tension in the strings, gravity, and air resistance. Determine q in terms of m, the mass of each ball; ℓ, the length of each string; d, the final distance between the balls; and any necessary constants. •••

50. Suppose you want to use a repulsive electric force to launch a rocket. (*a*) How much charge do you need to create enough repulsion to bring a 100,000-kg rocket to escape speed (11,200 m/s) while the rocket travels 100 mm? Assume that two identical charged objects, one mounted on the rocket, are separated by an average distance of 1.0 m for the whole launch. (*b*) If each elementary charge requires adding mass equivalent to the mass of one hydrogen atom to the rocket, how much additional mass would be required? •••

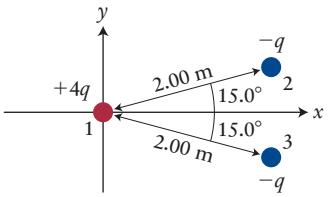
51. Two identical metal spheres carrying the same charge hang by insulating strings from the same point. The balls repel each other so that they hang in the air some distance apart. If you touch one of the spheres and let it go, what will the new equilibrium distance between the spheres be? •••

52. (*a*) Prove that the repulsive electric force between two charged particles is a maximum when each particle carries charge $q/2$, where q is the sum of the charges carried by the two particles. (*b*) Explain why this is *not* true for the attractive electric force between charged particles. •••

22.6 Forces exerted by distributions of charge carriers

53. Particle 1 carrying charge $+4q$ is fixed at the origin of a rectangular coordinate system, particle 2 carrying charge $-q$ is fixed 15.0° above the positive horizontal axis, and particle 3 carrying charge $-q$ is fixed 15.0° below the positive horizontal axis (Figure P22.53). Particles 2 and 3 are both 2.00 m from the origin. You must add two charged particles to the system in such a way that the vector sum of the electric forces exerted on particle 1 is zero. What quantity of which type of charge must each added particle carry, and at what locations must you place the particles? •

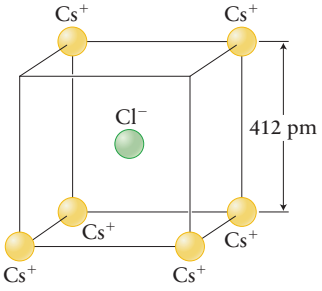
Figure P22.53



54. Two charges of equal magnitude, 2 μC, are separated by a 2 cm distance. Compare the electric force if the separation is increased to 6 cm. •

55. Cesium chloride (CsCl) is a crystalline salt that forms in a cubic lattice structure, which you can imagine as a cube with Cs^+ ions at the eight corners and a Cl^- ion at the center. The edge length of the cube is 412 pm. Suppose that at the edge of a crystal, two cesium atoms have been stripped from adjacent corners of a cube, as shown in Figure P22.55. What is the vector sum of the electric forces exerted on the chloride ion by the six Cs^+ ions remaining in the cube? •

Figure P22.55



56. Two particles carrying charges $q_1 = -4\ \mu C$ and $q_2 = 12\ \mu C$ are located 5 m apart. Choose a coordinate system where the first particle is at the origin, and the second is at $x = 5$ m. In this coordinate system, an electron sits at (2 m, −3 m). (*a*) Draw a diagram showing the vectors you need in order to determine the direction of the separate electric forces exerted on the electron by the particles. (*b*) Calculate the vector sum of these two forces. ••

57. Particle 1, carrying charge q, particle 2, carrying charge $2q$, and particle 3, carrying charge $-q$, are located at the vertices of an equilateral triangle of side ℓ in the xy plane. Particle 1 is located at $(-\ell\ /2,0)$, particle 2 is located at $(\ell\ /2,0)$, and particle 3 is located on the positive y axis. Find the direction of the electric force exerted on a fourth charged particle carrying charge q located at the midpoint between particles 2 and 3. ••

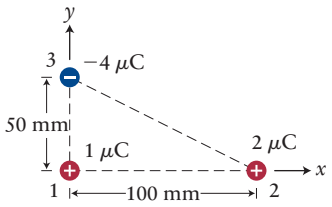
58. Four charged particles are located at the vertices of a square of side length 1m in the xy-plane. Particles 1, 2, and 3 ($q1 = 5.00\ \mu C$, $q2 = -4.00\ \mu C$, and $q3 = -3.00\ \mu C$) are located, respectively, at (0,0), (1 m, 0) and (0, 1 m). Determine the position of particle 4, carrying the same charge as particle 1, such that the total electric force on particle 1 is zero. ••

59. Four negatively charged particles each carry a charge q_n and form a square of side length d. A positively charged particle carrying a charge q_p is at the square's center. If the arrangement is in equilibrium (the vector sum of the electric forces on each particle is zero), what is the numerical value of the ratio q_p/q_n? ••

60. Particle 1 carrying 2.0 μC of charge is fixed at the origin of a rectangular coordinate system, particle 2 carrying $-1.0\ \mu C$ is fixed 45° above the positive horizontal axis 2.0 mm from the origin, and particle 3 carrying $-1.0\ \mu C$ is fixed 45° below the horizontal axis 2.0 mm from the origin. Where in the system should you place a particle 4 carrying $-2.0\ \mu C$ so that the vector sum of the electric forces exerted on particle 1 is zero? ••

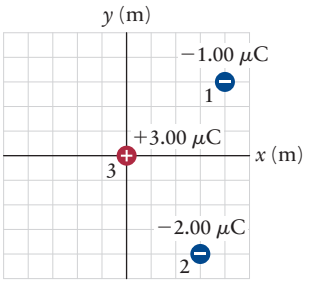
61. For the charge distribution in Figure P22.61, what is the ratio of the magnitude of the electric force exerted by particle 2 on particle 1 to the magnitude of the electric force exerted by particle 3 on particle 1? ••

Figure P22.61



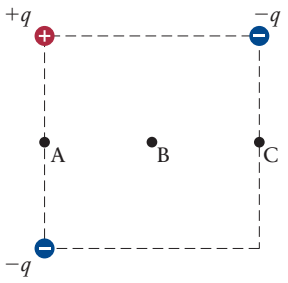
62. For the charge configuration shown in Figure P22.62, what is the vector sum of the electric forces exerted on particle 3? ••

Figure P22.62



63. Two particles, each carrying charge $-q$, are fixed at opposite corners of a square, and a third particle, carrying charge $+q$, is fixed at a third corner (Figure P22.63). (*a*) Draw an arrow representing the vector sum of the forces exerted on a particle carrying charge $+q$ as it is placed first at position A, then at position B, and then at position C. (*b*) Sketch the trajectory of this particle if it is released at each of these three positions. ●●

Figure P22.63



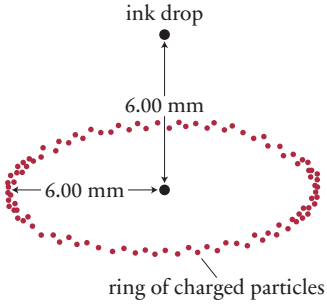
64. Two particles are located on the x axis of a Cartesian coordinate system. Particle 1 carries a charge of +2.0 nC and is at $x = -30$ mm, and particle 2 carries a charge of −2.0 nC and is at $x = 30$ mm. What are the magnitude and direction of the vector sum of the electric forces exerted by these two particles on particle 3, which carries a charge of $+5.0\ \mu$C and is located on the positive y axis 80 mm from the origin? ●●

65. Two identical charged dice having $+5\ \mu$C of charge, each having mass 3 g, stay in static equilibrium on the surface of a table with a separation of 12 m among them. Calculate the coefficient of static friction. ●●

66. Conducting spheres 1, 2, and 3 are placed at three corners of a square. Spheres 1 and 2 at opposing corners are very small and oppositely charged, and the magnitude of the charge on negatively charged sphere 1 is twice the magnitude of the charge on positively charged sphere 2. Sphere 3 is much larger than the other two and electrically neutral. (*a*) Draw the three spheres, showing the approximate distribution of charge on sphere 3. (*b*) Sketch a free-body diagram for sphere 2. ●●

67. You are working as an engineer designing ink-jet printers. Your boss has come across a printer design that involves a charged drop of ink experiencing a force exerted by 100 small, fixed charged particles. As shown in Figure P22.67, the particles are arranged in a ring of radius 6.00 mm, and each carries a charge of 1.00 nC. The ink drop is released at a point 6.00 mm away from the plane of the ring and along the central axis running perpendicular to the plane of the ring. Your boss gives up on calculating the initial force exerted on the charged drop, but you take one look at the setup and realize you can do the calculation in less than a minute. (*a*) Explain how you can do the calculation so quickly, even though it involves 100 separate forces. (*b*) Determine the magnitude and direction of the initial force exerted on a drop that carries a charge of 8.00 nC. (*c*) If the printer adds to each drop of ink a charge equal to 1.00 C per kilogram of ink mass, what is the initial acceleration of each drop? ●●●

Figure P22.67



68. A 0.160-kg hockey puck modified to hold $+50\ \mu$C of charge is placed at one end of an ice hockey rink that is 61 m long. At the other end of the rink is a device capable of generating a charge of −0.10 C, but only for a short time interval. The coefficient of kinetic friction between ice and puck is $\mu_k = 0.015$. The puck is given a tiny nudge (just enough to get it moving but at a speed of nearly 0, which means you can ignore static friction), and then the charge generator is turned on for 100 ms. (*a*) Is the puck able to reach the generator? (*b*) If so, what is the puck's speed when it arrives at the generator? (For both parts, assume the 100-ms time interval is short enough to allow you to assume that the distance between the puck and the charge generator is constant.) ●●●

69. Two particles 1 and 2, each carrying 3.0 nC of charge, are located 6 cm apart. Where along the symmetry axis between particles 1 and 2 should a particle 3 carrying a charge of +2.0 nC be located if it experiences an upward electric force of 1.2×10^{-5} N? ●●●

70. Figure P22.70 shows strips of transparent tape repelling each other. Use information given on the packaging of a roll of transparent tape to estimate what fraction of the atoms in either strip lose an electron in order to produce the repulsive effect. (Assume that the material of the tape is mostly carbon and hydrogen atoms present in a 1:2 ratio.) ●●●

Figure P22.70

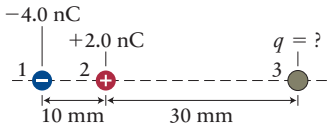


Additional Problems

71. Compare the ratio of electric force to gravitational force among two electrons and two protons, respectively, separated by 1 nm apart. Comment on the obtained different rations. ●

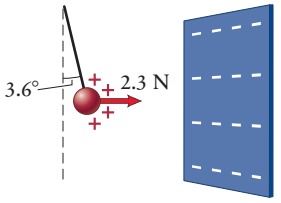
72. How many Coulombs of charge are contained due to the electrons present in 1 kg of liquid ammonia? •
73. A 10 μC negatively charged balloon having a total mass of 5 g that remains floated on top of a charged plastic board above 15 cm. Estimate the charge on the plastic board. •
74. Three identical metallic beads A, B, and C possess charges q_1, q_2, and q_3, respectively. All the three beads are affixed on a meter scale where A lies at 0, B lies at 60 cm, and C at 100 cm. In this configuration, if all the three beads attain electrostatic equilibrium, then find the charge ratios of q_2 to q_3. •
75. Particles A, B, and C, with charges 12.0 μC, 8.0 μC, and -6.0 μC, respectively, are evenly spaced along a line, in that order, with spacing 2.0 m. What is the magnitude and direction of the electric force exerted (*a*) on A, (*b*) on B, and (*c*) on C? •
76. A proton is composed of two up (*u*) quarks ($q = +2/3e$) and one down (*d*) quark ($q = -1/3e$). What is the electric force between two up quarks separated by 1×10^{-16} m inside a proton? ••
77. Three charged particles are arranged along a line as in Figure P22.77. If particle 2 experiences zero electric force, what is the charge on particle 3? ••

Figure P22.77



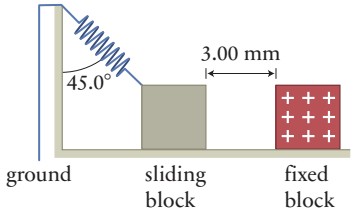
78. Four identical charged particles are constrained along an *x* axis. Identify one possible configuration of the particles that would leave one of them at rest at the origin when the others are fixed in place. No two charged particles can be at the same location. ••
79. (*a*) A particle carrying charge $+q$ is located at the center of a square with sides of length ℓ. Given four particles, two carrying charge $+q$ and two carrying charge $-q$, how should you place them at the four corners of the square, with one particle at each corner, such that the vector sum of the electric forces exerted on the center particle is zero? (*b*) If you leave the corner particles in the arrangement determined in part *a* but move the center particle to the midpoint of one side of the square, what are the magnitude and direction of the vector sum of the forces exerted on this particle? ••
80. A small charged sphere hanging from a thin string next to an oppositely charged plate experiences a 2.3-N attractive electric force (Figure P22.80). If the string makes an angle of 3.6° with the vertical, what is the mass of the sphere? ••

Figure P22.80



81. The accelerometer in a certain video game controller is constructed as shown in Figure P22.81. A block made of material that is an electrical conductor is grounded by a small stretched spring and is constrained to sliding along a rail. The net force on this block is zero. A nonconducting, positively charged block is fixed in position on the rail. When the magnitude of the electric force between the blocks is 0.100 N, the spring makes a 45.0° angle with the wall and is stretched 3.00 mm from its equilibrium position. What is the spring constant? Ignore friction. ••

Figure P22.81



82. An anti-hydrogen atom, A, is composed of one positron orbiting one antiproton just as a regular hydrogen atom is composed of one electron orbiting one proton. The difference is that each antiparticle has the usual mass but an opposite charge of its regular counterpart. Atom A is placed in a closed container with a neutron. What is the most likely charge of the system after the neutron decays? (Hint: Use charge conservation.) ••
83. All life on Earth is based on one of many two-polymer systems, each made up of a protein and a nucleic acid. In a cell, these polymers are, for the most part, charge-balanced, with the nucleic acid in any pair carrying a negative charge and the protein carrying an equal amount of positive charge. Consider a distribution of eight of these polymers, where the eight proteins sit together near the center of the cell and the nucleic acids are spread out along the circumference of a circle with the radius equal to one-quarter of the cell diameter, centered on the proteins. If seven of the eight nucleic acids are at $\pi/4$, $\pi/2$, $3\pi/4$, π, $3\pi/2$, $7\pi/4$, and 2π rad, where on the circle must you place the eighth acid to give the system a balanced charge distribution? (Hint: Begin by making a sketch.) ••
84. In a simplistic model of the hydrogen atom, the electron orbits the proton in a circle of radius 53 pm. What is the orbital period of the electron, in seconds, if the force responsible for the proton-electron attraction is (*a*) gravitational and (*b*) electric? ••

85. (*a*) About how many electrons are there in a copper penny, which has a mass of about 0.003 kg? (*b*) If you could isolate these electrons, how much charge would you have? (*c*) Estimate the magnitude of the electric force it would require to bring one more electron to within 1.0 nm of these electrons, assuming they are combined to form one small charged object. (*d*) Based on this force magnitude, do you think it would be possible to combine this amount of negative charge? ●●

86. Earth exerts an electric force on small charged objects located near its surface. The effect can be modeled by assuming Earth is a particle at the center of Earth carrying a charge of -6.76×10^5 C. (*a*) What is the magnitude of the electric force exerted by Earth on an electron at its surface? Is this force attractive or repulsive? (*b*) How does this force magnitude compare with the magnitude of the gravitational force exerted by Earth on the electron? (*c*) How much negative charge would you have to put on a penny in order for the repulsive electric force exerted by Earth on the penny to cancel the attractive gravitational force exerted by Earth on the penny? (*d*) How many electrons would you have to add to the penny to achieve this? ●●

87. Four particles, each carrying a charge of magnitude 3.0 nC, are located at the corners of a square of side 50 mm. Two particles have positive charges and two have negative charges. (*a*) Determine the position of each particle so that the vector sum of the electric forces on each particle is directed toward the center of the square. (*b*) Determine the magnitude of the vector sum of the electric forces exerted on each particle. ●●

88. A small sphere 1 carrying charge q is fixed a few meters above the Moon's surface next to a spaceship. An astronaut holds a charged sphere 2 a couple of meters above sphere 1, almost but not quite directly over sphere 1. Sketch the trajectory for sphere 2 when the astronaut drops it if the charge on sphere 2 is (*a*) q, (*b*) $9q$, and (*c*) $-q$. ●●●

89. Someone has challenged you in a friendly bet that, given a set of six charged spheres with three carrying positive charge and three carrying negative charge, you can't place any three of them in a line (a board with a groove in it serves the purpose) and have them be in translational equilibrium. You determine which of the three spheres are the positively charged ones and call them 1, 2, and 3. By comparing the electric force exerted by sphere 1 on sphere 3 with the electric force exerted by sphere 2 on sphere 3, you rank 1 as smaller than 2, placing them in the groove several hundred millimeters apart, temporarily blocking them from moving. You can release them only when the third sphere is in place, but there are so many things to consider! ●●● CR

90. While assembling part of an electric generator, you realize that two pieces were not properly labeled. One piece is a nonconducting block, and the other is a conducting block, but without labels you don't know which is which. Because you had counted on the parts being properly labeled, you do not have a multimeter, batteries, or other standard electrical tools. All you have is your wool sweater, some nonconducting styrofoam "peanuts" used in packaging, and some string. You are about to go for help when you realize that your knowledge of static electricity should be enough. ●●● CR

91. Working on a system to collect solar energy, you are trying to move some charged particles through the open interior of a uniformly charged ring. A colleague vaguely remembers Newton's and Priestley's argument that no electrostatic force is exerted on a charged particle inside a uniformly charged sphere. He suggests that the same must be true for a charged ring, and so you can send the charged particles through any part of the ring interior without the particles feeling an electric force and being deflected by it. You are skeptical, wondering if there might at least be a restriction on the path over which you send the particles such that they experience no electric force. ●●● CR



23

The Electric Field

23.1 The field model
23.2 Electric field diagrams
23.3 Superposition of electric fields
23.4 Electric fields and forces

23.5 Electric field of a charged particle
23.6 Dipole field
23.7 Electric fields of continuous charge distributions
23.8 Dipoles in electric fields

CONCEPTS

QUANTITATIVE TOOLS

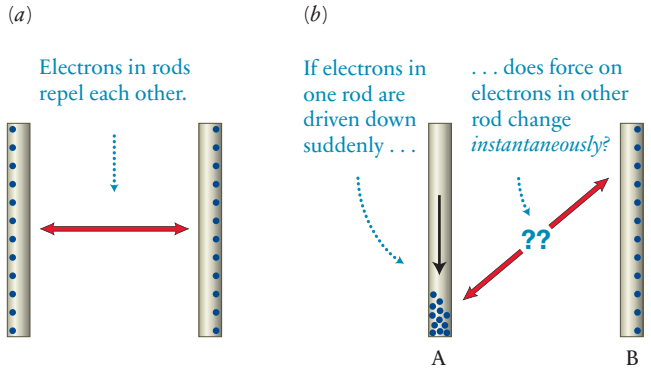
In this chapter we revisit an issue we discussed briefly in Chapter 7: the long-range nature of electric and gravitational interactions. How does one charged object "reach out" and affect another charged object? What are the invisible "springs" that pull us—and everything around us—toward Earth's surface? You can describe these long-range interactions by saying that every charged object and every object that has mass has a "sphere of influence" surrounding it. The modern word for this sphere of influence is *field*. Fields are not imaginary. That sensation you felt when, while reading the beginning of Chapter 22, you held a piece of plastic food wrap near your face was the sensation of a field created by the charged particles in the wrap. The closer the wrap is to your skin, the stronger the sensation.

The concept of field is important for two reasons. First, it is impossible to describe the interaction between moving charged particles without it. Second, as you will soon see, it is often easier to deal with fields than with distributions of charge because frequently more is known about fields than about the way charge is distributed.

23.1 The field model

Newton's law of gravity and Coulomb's law describing the electric force between charged particles successfully account for the magnitudes of the gravitational and electric forces between stationary objects. However, they do not address the fundamental puzzle of how objects separated in space can interact without any mediator of the interaction (such an interaction is called *action at a distance*). Worse, they share a fundamental flaw: Both imply that the action of one object on another is instantaneous everywhere throughout space. Consider, for example, the two metal rods in **Figure 23.1**. Even if both are electrically neutral, their electrons interact. Suppose you quickly drive the electrons in rod A down to the bottom, as in Figure 23.1*b*. According to Coulomb's law, doing this will instantly change the force exerted by the electrons in A on those in B, regardless of how far apart the rods are. This means that it would be possible—in principle—to be standing at one position in space and instantly detect a change that occurs at some far distant position.

Figure 23.1 Newton's and Coulomb's laws imply that forces are exerted instantaneously across a distance—but experiments show that they are not.



The idea that an object can directly and instantly influence another object regardless of their separation was troubling in the 19th century but became untenable in the early 20th century when it was demonstrated experimentally that the interaction between charged objects is not instantaneous. The principle illustrated in Figure 23.1, for example, is what makes possible the transmission of radio signals from a transmitting antenna (rod A) to a receiving antenna (rod B). We know from many experiments that such transmission is not instantaneous. As just one example, a radio signal takes about 2.5 s to travel from Earth's surface to the surface of the Moon. In other words, the picture conveyed by Newton's law of gravity and Coulomb's law—that one object directly and instantly affects other objects regardless of the distance between them—cannot be correct.

Instead we must adopt another model of long-range interactions, a model in which interactions take place through the intermediary of an **interaction field** (or simply a **field**). In the field model, an interacting object fills the space around itself with a field. When an object A is placed in the field of an object B, A can "feel" the presence of B's field. Instead of the two objects interacting directly as illustrated in **Figure 23.2*a***, it is the field created by each object that acts on the other object (Figure 23.2*b*). The stronger the field, the greater the magnitude of the force resulting from the interaction.

In Figure 23.1, the electrons in rod B feel the field set up by the electrons in rod A. When the electrons in A accelerate, their motion causes a disturbance in A's field, and this disturbance propagates outward through space like the ripples on the surface of a pond. Only when these ripples in the field reach rod B can the motion of A's electrons be detected by those in rod B. In Chapter 30 we shall study the propagation of disturbances in fields due to accelerating charge carriers. For now, we shall concentrate on fields created by stationary objects.

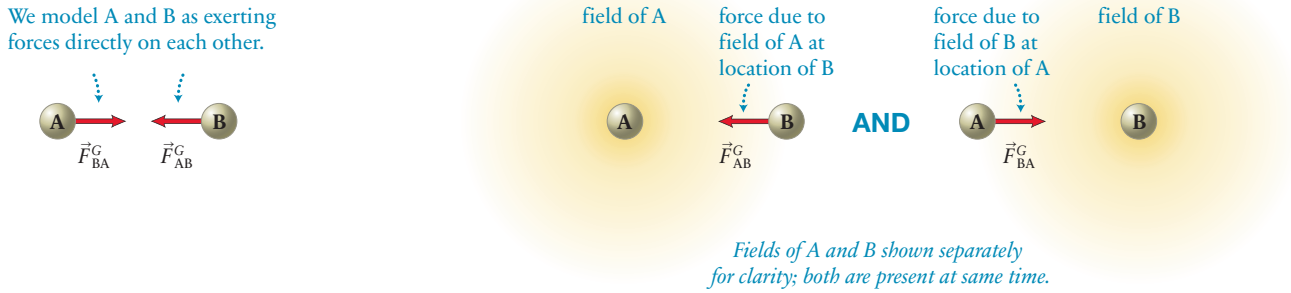
The field model applies equally well to gravitational and electric interactions, with each interaction having its own type of field. The space around any object that has mass is filled with a *gravitational field*, and the space around any electrically charged object is filled with an *electric field*. Gravitational fields exert forces on objects that have mass, and electric fields exert forces on objects that either carry a charge or can be polarized. Let's begin by developing the concept of gravitational field.

Before we attempt to obtain a physical quantity we can use to describe any gravitational field, we should note a number of things. First, for any object A located in a gravitational field created by an object S (S is called the *source* of the field), the magnitude of the field felt by A depends only on the properties of S and on the position of A relative to S; the field magnitude does not depend in any way on the properties of A.

CONCEPTS

Figure 23.2 The field model for interaction at a distance.

(*a*) Model of direct interaction at a distance (*b*) Field model of interaction at a distance

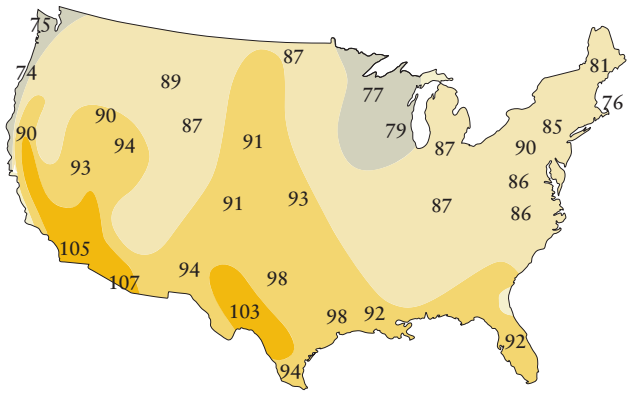


Second, the field of an object is always there, even when the object is not interacting with anything else. A field therefore must be represented by a set of numerical values that cover all the space outside the field source, with every point of this space having a different numerical value. One field representation you are already familiar with is that of a "temperature field," where the temperature across the surface of a region has a specific value at each location (**Figure 23.3**). Third, for stationary objects the field model must give the same forces as Newton's law of gravity and Coulomb's law. In particular, the field model must still yield forces between the two objects that are equal in magnitude and opposite in direction. It is not immediately obvious that the field model preserves this symmetry in the forces because a field is not something shared by two interacting objects—each object has its own field.

For example, in the gravitational interaction between a ball and Earth, the gravitational force exerted by Earth on the ball is due to Earth's gravitational field and the gravitational force exerted by the ball on Earth is due to the ball's field. These two gravitational fields are very different from each other: Earth's field pulls strongly on, say, a paper clip, while the effect of the ball's field on that paper clip is unmeasurably small. As you will see shortly, however, the symmetry of the interaction is preserved despite the asymmetry in the fields (see Checkpoint 23.3).

What physical quantity can we use to describe the gravitational field of an object? How about the gravitational force exerted by the object? Let's examine this possibility using Earth as our object. As we saw in Section 8.8, the magnitude of the gravitational force exerted by Earth on an object of mass m near Earth's surface is $F^G_{Eo} = mg$. This force is not a good quantity for describing Earth's field because the force depends not only on the source—Earth (which determines g)—but also on the mass m of the object placed in the field. As illustrated in **Figure 23.4**, two objects that have different masses m_1 and m_2 but are placed at the same height above Earth's surface are subject to different gravitational forces m_1g and m_2g. The quantity $g = F^G_E/m$, however—the gravitational force per unit of mass—is the same for any object.* This quantity is determined solely by the properties of Earth and is independent of those of any object that experiences a gravitational force exerted by Earth.

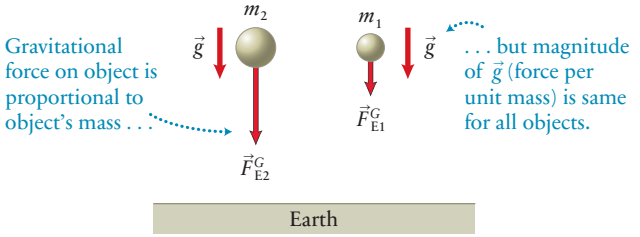
Figure 23.3 The temperature across a region is specified by a set of values, with a specific temperature value for every position in that region. Such a set of values is called a *field*.



23.1 Two objects 1 and 2, of mass m_1 and m_2, are released from rest far from Earth, at a location where the magnitude of the acceleration due to gravity is much less than $g = 9.8\ \text{m/s}^2$. (*a*) What is the ratio F^G_{E1}/F^G_{E2}? (*b*) For these two objects, is the magnitude of the gravitational force exerted by Earth per unit mass independent of the properties of the objects?

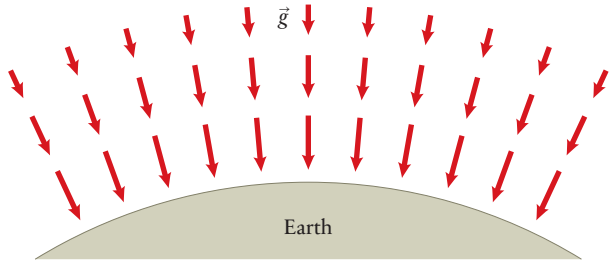
CONCEPTS

Figure 23.4 Comparison between gravitational force and gravitational acceleration on objects of different mass at the same distance from Earth.



*This is so because mass and inertia are equivalent (see Section 13.1), and so the gravitational force per unit of mass is equal to gravitational force divided by inertia, which is gravitational acceleration.

Figure 23.5 Vector field diagram for the gravitational field in a region near Earth.



We can use the gravitational force per unit mass exerted by an object as a measure of the magnitude of the object's gravitational field. For example, near Earth's surface, because $g = 9.8\ \text{m/s}^2$, the magnitude of Earth's gravitational field is 9.8 N/kg; near the surface of the Moon, where $g_{\text{moon}} = 1.6\ \text{m/s}^2$, the magnitude of the Moon's gravitational field is 1.6 N/kg. (Remember that $1\ \text{N/kg} = 1\ (\text{kg}\cdot\text{m/s}^2)/\text{kg} = 1\ \text{m/s}^2$.)

At any given location in the space surrounding a source object S, the magnitude of the gravitational field created by S is the magnitude of the gravitational force exerted on an object B placed at that location divided by the mass of B.

Unlike the temperature field in Figure 23.3, which is a *scalar field*, the gravitational field is a *vector field:* At every position, it has both a magnitude and a direction. **Figure 23.5**, for example, shows a **vector field diagram** representing the gravitational field near Earth. You can determine the magnitude and direction of this field in the space surrounding Earth using a **test particle** (an idealized particle whose mass is small enough that its presence does not perturb the object whose gravitational field we are measuring). Measure, at each location, the gravitational force exerted by Earth on the test particle, and then divide that force by the mass of the test particle to obtain the direction and the magnitude of the gravitational field at that location. As you can see from Figure 23.5, Earth's gravitational field, which can be represented at each position by a vector $\vec{g}$, always points toward the center of Earth, and its magnitude decreases with increasing distance away from Earth. Near Earth's surface, the magnitude of these vectors is $g = 9.8$ N/kg.

23.2 A communications satellite orbits 1.4×10^7 m from Earth's center, at a location where the magnitude of Earth's gravitational field is 2.0 N/kg. (*a*) If the mass of the satellite is $m_s = 2000$ kg, what is the magnitude of $\vec{F}^G_{\text{Es}}$? (*b*) If you place a 0.20-kg ball at the satellite's location, what is the magnitude of $\vec{F}^G_{\text{Eb}}$?

Checkpoint 23.2 illustrates that if you know the gravitational field at a certain position, you can easily calculate the gravitational force exerted by the source of that field on any object at that position by taking the product of the magnitude of the gravitational field at the location of the object and the mass of the object.

23.3 (*a*) Is the magnitude of the gravitational force exerted by Earth on a ball greater than, equal to, or smaller than the magnitude of the gravitational force exerted by the ball on Earth? (*b*) Is the magnitude of Earth's gravitational field at the position of the ball greater than, equal to, or smaller than the magnitude of the gravitational field of the ball at a distance equal to Earth's radius? (*c*) Explain how the answers to parts *a* and *b* can both be correct.

23.2 Electric field diagrams

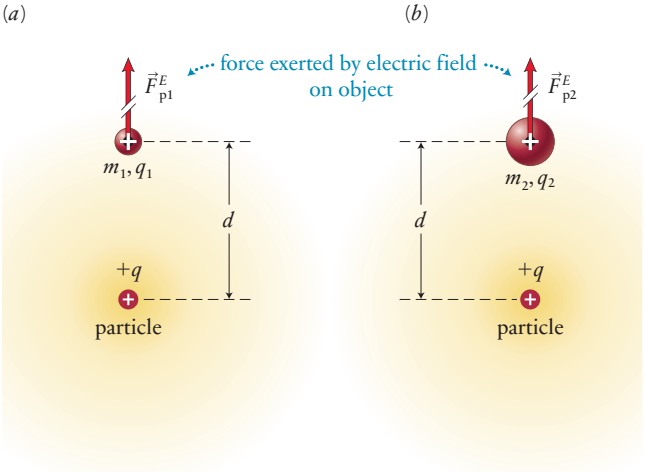
Let us now apply the same ideas to electric interactions. **Figure 23.6***a* shows object 1 of mass m_1 and charge q_1 a distance d from a particle that carries a charge $+q$. What is the electric field $\vec{E}$ created by the particle at the position of object 1? Before answering this question, answer the next checkpoint, which concerns the interactions of the particle with object 1 and with an object 2 of mass m_2 and charge q_2 (Figure 23.6*b*; $m_2 \neq m_1$ and $q_2 \neq q_1$).

23.4 (*a*) Are the electric forces $\vec{F}^E_{\text{p1}}$ and $\vec{F}^E_{\text{p2}}$ in Figure 23.6 equal? (*b*) What does the quantity $\vec{F}^E_{\text{pi}}/m_i$ represent? (*c*) Is this quantity the same for objects 1 and 2? If not, what quantity is the same for both of these objects?

As Checkpoint 23.4 makes clear, the quantity $\vec{F}^E_{\text{pi}}/q_i$—the electric force per unit charge—is determined entirely by the source of the electric field and is independent of the object on which the field exerts a force. So, in analogy to the gravitational field, we can say:

At any given location in the space surrounding a source object S, the electric field created by S is the electric force exerted on a charged test particle placed at that location divided by the charge of the test particle: $\vec{E}_{\text{S}} \equiv \vec{F}^E_{\text{St}}/q_{\text{t}}$.

Figure 23.6 Electric force exerted on two objects of different inertia m and charge q by the electric fields created by two identical charged particles.



CONCEPTS

Like gravitational fields, electric fields are vector fields. There is one difference between the two types of fields, however. Electric interactions can be either repulsive or attractive, and so the direction of the electric force—and hence the direction of the electric field—depends on the sign of the charge. Our rule is that:

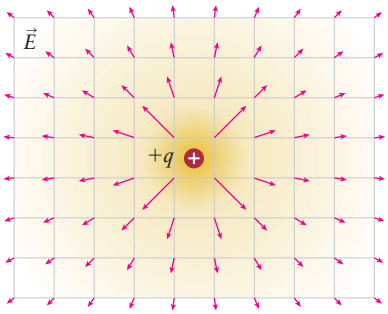
The direction of the electric field at a given location is the same as the direction of the electric force exerted on a positively charged object at that location.

23.5 (*a*) If the particle in Figure 23.6 carries a negative charge $q < 0$ and q_1 and q_2 are positive, what are the directions of $\vec{F}^E_{p1}$ and $\vec{F}^E_{p2}$? (*b*) Does the electric field created by the particle point toward or away from the particle? (*c*) If q and q_2 are negative, what are the direction of $\vec{F}^E_{p2}$ and the direction of the electric field created by the particle at the location of object 2? (*d*) If q is positive, does the electric field created by the particle point toward or away from the particle? (*e*) How does the magnitude of the electric field created by a particle that carries a charge $+q$ ($q > 0$) compare with the magnitude of the electric field created by a particle that carries a charge $-q$ of identical magnitude at a distance d from each particle?

Figure 23.7 shows the vector field diagrams for the electric fields of particles that carry positive and negative charges. Because it is impossible to draw electric field vectors at *all* locations, the diagrams show vectors at only certain positions; from these representative vectors you can get an idea of how the electric field looks as a whole. In addition, the drawing is limited to two dimensions, but you should visualize the electric field as spreading out in all three dimensions.

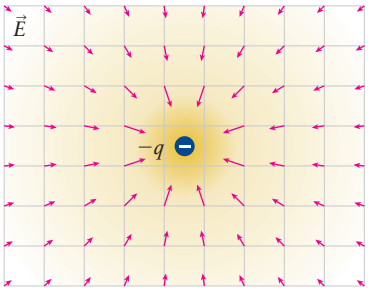
Figure 23.7 Vector field diagrams for positively and negatively charged particles. The lengths of the vectors show that the electric field magnitude decreases with increasing distance from the source.

(*a*) Electric field of positively charged particle



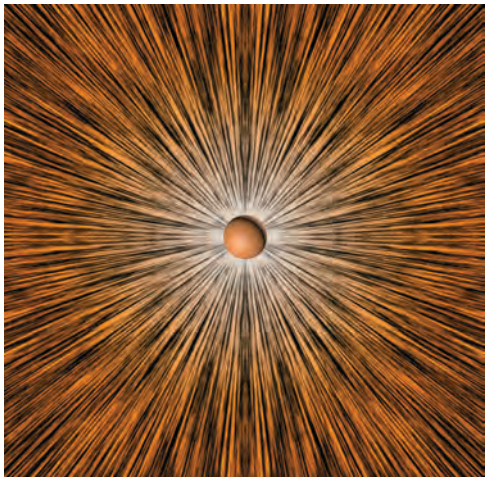
Electric field is directed away from positive source . . .

(*b*) Electric field of negatively charged particle



. . . and toward negative source.

Figure 23.8 Electric field pattern created by a small charged object in a solution that contains plastic fibers. The fibers align with the direction of the electric field created by the charged object.



Electric fields can be made visible by putting charged objects in a (nonconducting) liquid that contains small uncharged plastic fibers or grass seed. Each fiber aligns itself in the direction of the electric field at the fiber's location (**Figure 23.8**).

23.6 If you know the electric field $\vec{E}$ at some location, how can you determine the magnitude and direction of the electric force exerted by that field on an object carrying a charge q and placed at that location?

CONCEPTS

23.3 Superposition of electric fields

The concept of electric field becomes especially useful when we consider the combined electric field that results from more than one charged object. Suppose we are interested in the electric field created by two particles that carry charges of equal magnitude but opposite sign. To determine the electric field created by the particles at a point P, we place a test particle* carrying a positive charge q_t at P and measure the vector sum of the forces exerted on it by the two charged source particles (**Figure 23.9** on the next page). The electric field at P is then equal to this vector sum divided by q_t.

*When measuring electric fields, we assume that the charge q_t on the test particle is so small that the particle does not perturb the particles or objects that generate the electric field we are measuring.

Figure 23.9 The electric field due to multiple charged objects (here, a pair of charged particles) is the vector sum of the fields created by the individual objects.

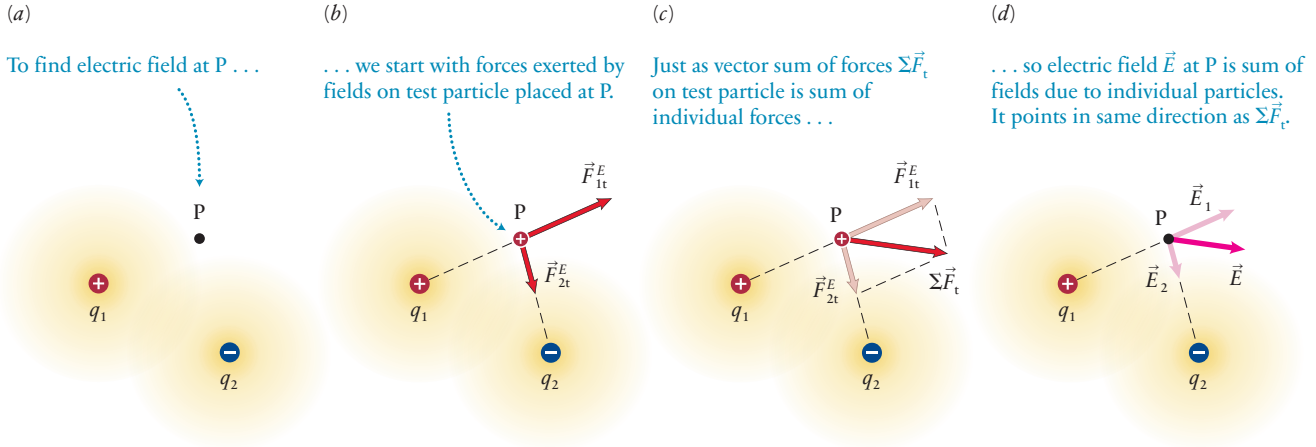


Figure 23.9 illustrates the **superposition of electric fields:**

The combined electric field created by a collection of charged objects is equal to the vector sum of the electric fields created by the individual objects.

The superposition principle holds regardless of the number of sources. Because of the vectorial nature of the electric interaction, electric forces add vectorially (Eq. 22.9). Consequently the electric field at P is equal to $(\vec{F}^E_{1t} + \vec{F}^E_{2t})/q_t = \vec{F}^E_{1t}/q_t + \vec{F}^E_{2t}/q_t$, which is the vector sum of the electric fields created by the two sources individually.

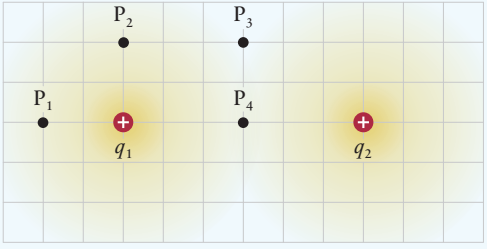
The only caveat is the one I pointed out in Figure 22.30. When we deal with conductors, the distribution of charge on the individual conductors in isolation might be different from what it is when the conductors are placed close together.

CONCEPTS

EXERCISE 23.1 Electric field of two positively charged particles

Consider two identical particles 1 and 2 carrying charges $q_1 = q_2 > 0$ (**Figure 23.10**). What is the direction of the combined electric field at points P_1 through P_4?

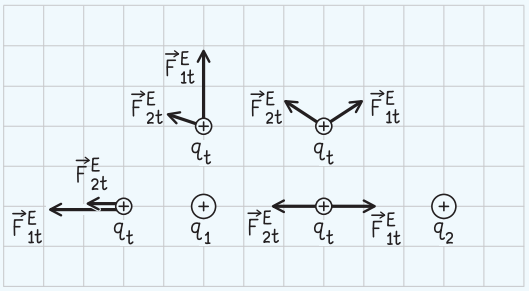
Figure 23.10 Exercise 23.1.



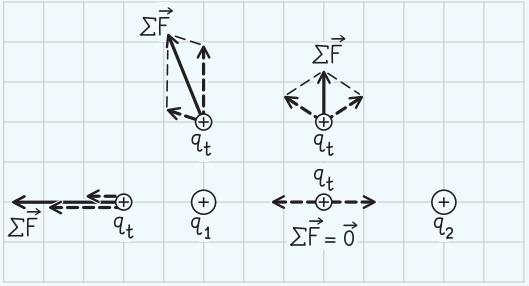
SOLUTION I place a positively charged test particle at each location and determine the vector sum of the (repulsive) forces exerted by 1 and 2 on each test particle. Because $\vec{E} = \Sigma\vec{F}/q_{test}$, the direction of $\vec{E}$ is the same as the direction of $\Sigma\vec{F}$ (**Figure 23.11**).

Figure 23.11

(a) Electric forces on test particles



(b) Vector sum of forces on each test particle



(c) Electric field at each tested point

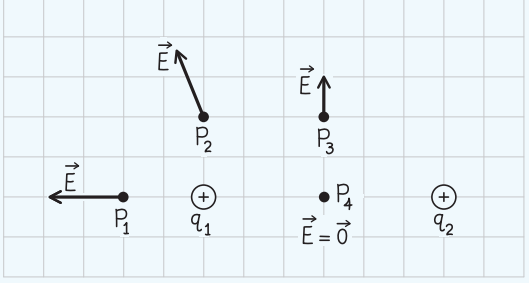
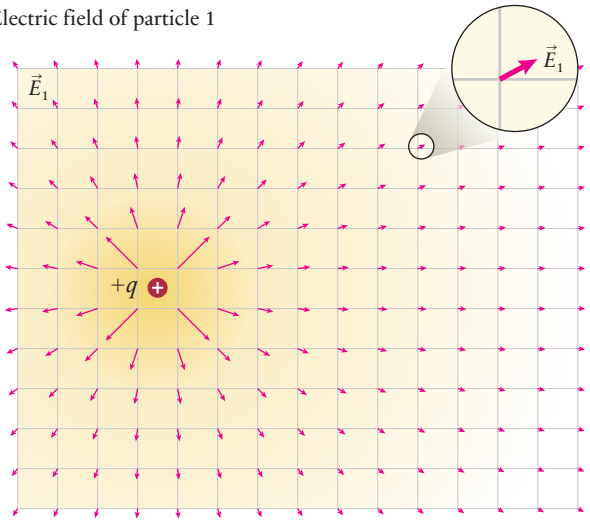
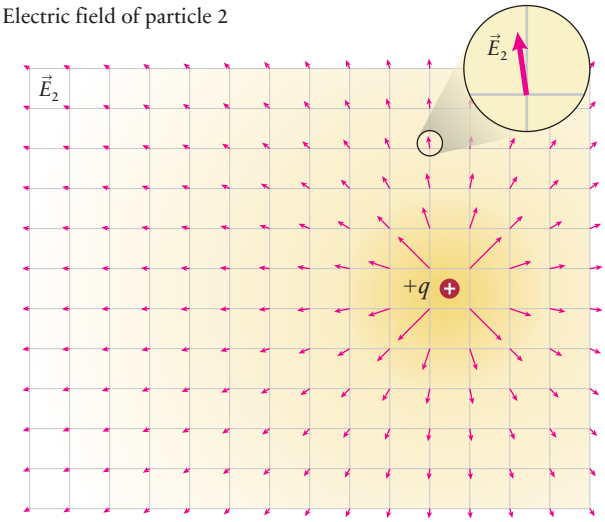


Figure 23.12 Vector field diagrams showing the superposition of the electric fields of the two charged particles of Figure 23.10.

(*a*) Electric field of particle 1



(*b*) Electric field of particle 2



(*c*) Electric field of both particles

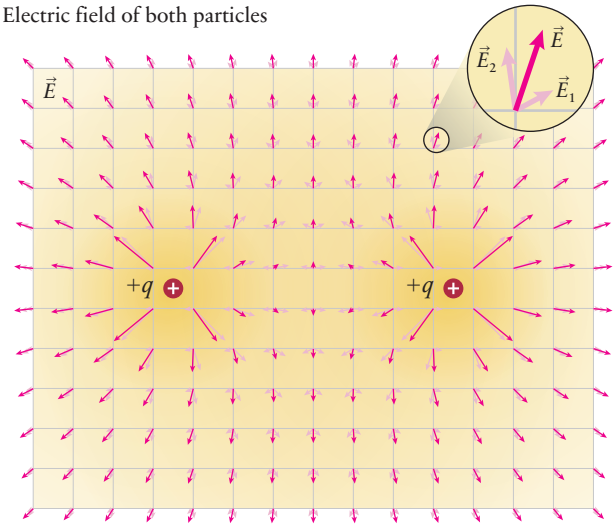
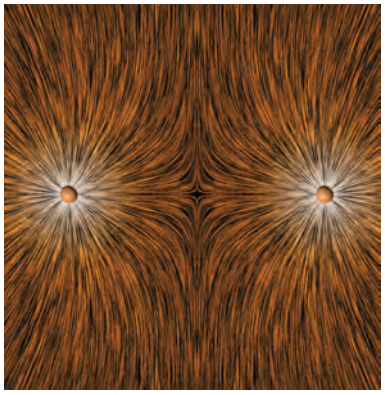


Figure 23.13 Pattern created by two identical charged particles in a liquid containing plastic fibers. Compare with the vector field diagram in Figure 23.12*c*.

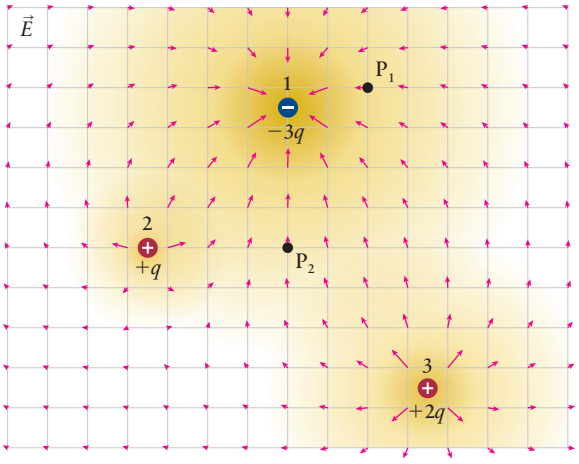


23.7 (*a*) If the charge on particle 2 in Exercise 23.1 is doubled so that $q_2 = 2q_1$, what happens to the direction of the electric field at points P_1 through P_4? (*b*) If the charge on particle 2 is negative so that $q_2 = -q_1$, what is the direction of the electric field at points P_1 through P_4?

Figure 23.11 provides a limited view of the electric field created by the two particles. A more complete view is given in **Figure 23.12**. This diagram is obtained by vectorially adding, for each grid point, the electric field vectors for the individual particles. Note how the pattern of vectors resembles the pattern created by two identically charged particles in a solution of plastic fibers (**Figure 23.13**).

Using the superposition principle, we can determine the electric field produced by any system of charged particles. **Figure 23.14**, for example, shows a vector diagram for the electric field generated by three charged particles. Because every charged object is made up of charged particles—electrons and protons—we can determine the electric field of any object at any position in space. For a real object, the calculation might

Figure 23.14 Vector field diagram of the electric field created by three charged objects.



CONCEPTS

be very tedious or even intractable because of the large number of charged particles, but the basic principle is as given above.

23.8 (*a*) In Figure 23.14, what is the direction of the force $\sum\vec{F}^E$ exerted on a particle carrying a charge $+q$ and placed at P_1? (*b*) How does $\sum\vec{F}^E$ change when the particle at P_1 carries a charge $+3q$? (*c*) Do the magnitude and direction of $\sum\vec{F}^E$ on a charged particle at P_1 change if the $+2q$ charge on object 3 is halved? (*d*) What is the direction of $\sum\vec{F}^E$ exerted on a particle carrying a charge $-q$ and placed at point P_2?

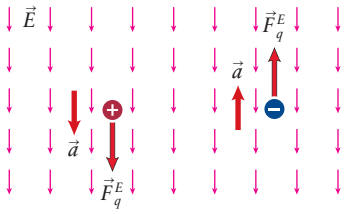
23.4 Electric fields and forces

Before developing additional techniques for determining the electric fields created by systems of charged particles, let us consider this question: What are the forces exerted by an electric field on charged or polarized objects? One of the advantages of working with electric fields is that, for any system of charged particles, once we know the electric field the system creates at some point P in space, we can determine the force exerted by the system on any other charged particle placed at P without worrying about any of the individual source particles in the system anymore.* In Chapter 22 we used the action-at-a-distance model to discuss the forces exerted by charged objects either on other charged objects or on polarized objects. Now we can use the field model to do the same thing. For stationary charged particles, both methods must yield the same result.

When we study the forces exerted by electric fields, it is useful to distinguish between uniform and nonuniform. In a *uniform electric field*, the direction and magnitude of the electric field are the same everywhere. No electric field is ever uniform throughout all space, but as we shall see in Section 23.7 it is possible to create regions of space where the electric field is uniform. In a *nonuniform electric field*, the direction and magnitude of the electric field vary from position to position. (All the electric fields we have considered so far are nonuniform.)

Let us first consider what happens to a charged particle placed in a uniform electric field where the only interaction is electrostatic. Because the electric field is defined as the electric force per unit of charge, the force $\vec{F}_p^E$ exerted by an electric field $\vec{E}$ on a particle carrying a charge q is $\vec{F}_p^E = q\vec{E}$.† Because $\vec{E}$ is the same everywhere, the force $\vec{F}_p^E$ exerted on the particle is constant and so it undergoes a constant acceleration $\vec{a} = \vec{F}_p^E/m = q\vec{E}/m = (q/m)\vec{E}$, where m is the particle's mass.

Figure 23.15 Forces exerted by a uniform electric field on a positively and a negatively charged particle.



A charged particle placed in a uniform electric field undergoes constant acceleration.

If the particle carries a positive charge, $q > 0$, $\vec{F}_p^E$ and $\vec{a}$ point in the same direction as $\vec{E}$. If $q < 0$, $\vec{F}_p^E$ and $\vec{a}$ point in the direction opposite the direction of the electric field (Figure 23.15).

Note from $\vec{a} = (q/m)\vec{E}$ that the magnitude of the acceleration depends on the magnitude of the electric field and on the charge-to-mass ratio q/m of the particle. A large charge q causes a greater force to be exerted on the particle and therefore a greater acceleration; a larger mass m means the particle has greater inertia and therefore the acceleration is smaller.

Because we have already studied motion with constant acceleration, we can apply our knowledge to the motion of charged particles in a uniform electric field. In general, the trajectory of these particles is parabolic, like the trajectory of a projectile fired near Earth's surface, where the gravitational field can be considered uniform over a limited area. In the special case where the initial velocity of a charged particle is parallel to the direction of the electric field, the trajectory is a straight line, like the vertical fall of an object released from rest. The main difference between the motion of projectiles near Earth's surface and the motion of charged particles in an electric field is that Earth's gravitational field is always directed vertically downward, whereas the electric field can be in any direction.

EXAMPLE 23.2 Charged particle trajectories

Four charged particles are fired with a horizontal initial velocity $\vec{v}_i$ into a uniform electric field that is directed vertically downward. The effect of gravity is negligible. The particles have the following charges and masses: particle 1 $(+q, m)$; 2 $(+q, 2m)$; 3 $(+2q, 2m)$; 4 $(-q, m)$. Sketch the four trajectories.

1 GETTING STARTED Because the electric field direction is vertically down, the three positively charged particles experience a downward force and the negatively charged particle experiences an upward force. The magnitude of this force doesn't change as the particles move through the

*This procedure is valid only if the presence of the charged particle at P does not alter the way charge is distributed over the system. We shall refer to the particle or system of particles that creates the electric field at P as being "fixed."

†When dealing with forces exerted by fields, we drop the subscript representing the object that exerts the force (the "by" subscript) because the field is due to *all* objects other than the object on which the force is exerted. The superscript E reminds us that we are dealing with the force exerted by an electric field.

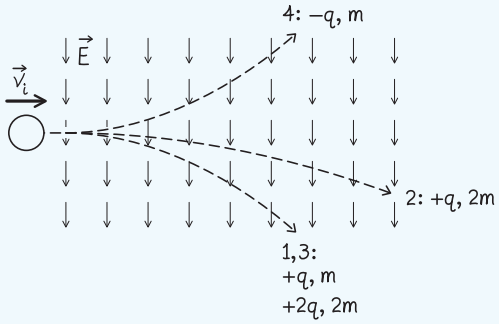
CONCEPTS

electric field because both the field magnitude and the charges on the particles are constant.

2 DEVISE PLAN Because the force exerted on each particle is constant, the particles experience constant accelerations. Because the direction of the force is perpendicular to the direction of the particles' initial motion, they all have a parabolic trajectory. The positively charged particles have a constant downward acceleration; the negatively charged particle has a constant upward acceleration. Because $\vec{a} = (q/m)\vec{E}$, the acceleration magnitude is greatest when q is large and/or m is small.

3 EXECUTE PLAN I draw trajectories that curve down for 1, 2, and 3 and up for 4 (**Figure 23.16**). The magnitude of the electric force exerted on particle 2 is the same as that exerted on particle 1, but the acceleration of particle 2 is smaller because this particle has the greater mass. I indicate this difference in acceleration by making trajectory 1 more curved than trajectory 2. The magnitude of the electric force exerted on particle 3 is twice as great as that exerted on particle 1, but 3's mass is also twice as great, and so the two particles have the same charge-to-mass ratio and therefore the same acceleration and trajectory. The magnitude of the electric force exerted on particle 4 is the same as that exerted on particle 1 but points in the opposite direction, and so trajectories 1 and 4 are identical in shape but curve in opposite directions. ✔

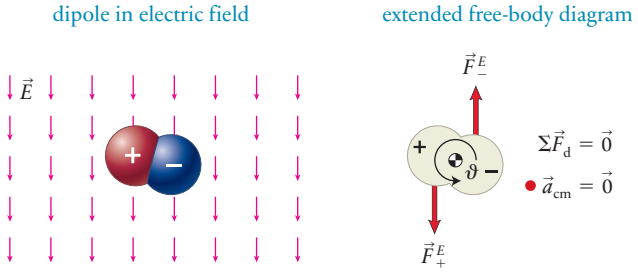
Figure 23.16



4 EVALUATE RESULT My sketch indicates that particles with increasingly positive charge-to-mass ratios curve increasingly downward. Conversely, particles with increasingly negative charge-to-mass ratios curve increasingly upward. This is what I expect because a particle's deflection is a function of both its charge, which determines the magnitude of the force exerted by the electric field on it (greater charge, greater deflection), and its mass, which relates the particle's acceleration to the force exerted on it (greater mass, smaller deflection).

23.9 A water droplet carrying a positive charge is released from rest in a uniform horizontal electric field near Earth's surface. The horizontal electric force is comparable in magnitude to the gravitational force exerted by Earth. Describe the droplet's trajectory.

Figure 23.17 Extended free-body diagram for a permanent dipole placed in a uniform electric field.



In a nonuniform electric field, the force exerted on a charged particle varies from one position to another, so we cannot easily specify the particle's trajectory without knowing more about the electric field. As in a uniform electric field, however:

A positively charged particle placed in a nonuniform electric field has an acceleration in the same direction as the electric field; a negatively charged particle placed in such a field has an acceleration in the opposite direction.

In Chapter 22 we found that charged objects can polarize electrically neutral objects by separating the centers of positive and negative charge in the latter. The resulting configuration of charge—equal amounts of positive and negative charge separated by a small distance—is called an **electric dipole** or simply **dipole.** Many molecules, such as water molecules, are *permanent dipoles;* that is to say, the centers of positive and negative charge are kept separated by some internal mechanism. **Figure 23.17** illustrates the forces exerted on a permanent electric dipole in a uniform electric field. Because the electric field is uniform and the magnitude of the charge on the positive end of the dipole is equal to the magnitude of the charge on the negative end, the forces exerted on the two ends are equal in magnitude but opposite in direction, making their vector sum zero. However, the forces exerted on the two ends cause a torque (see Chapter 12).

23.10 (*a*) What effect does the torque caused by the electric field have on the electric dipole in Figure 23.17? (*b*) Is the torque the same for every orientation of the molecule?

The orientation of an electric dipole can be characterized by a vector, the **dipole moment,** that, by definition, points from the center of negative charge to

CONCEPTS

the center of positive charge, as shown in **Figure 23.18**. As Checkpoint 23.10 illustrates, the electric forces create a torque on the dipole that tends to align the dipole moment with the electric field.

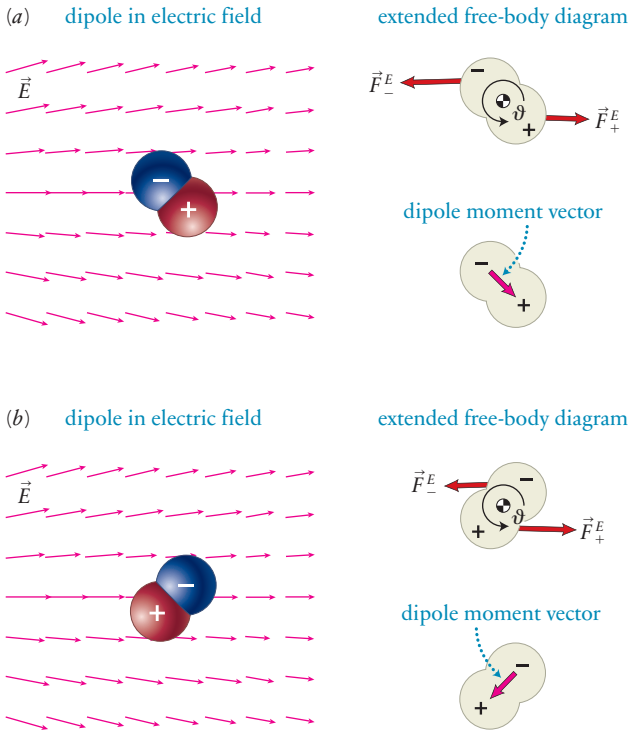
In a nonuniform electric field, the situation is more complicated because the two ends of the dipole are now subject to forces that have different magnitudes as well as different directions. Consider, for example, the nonuniform electric field in Figure 23.18*a*, which is due to a positively charged particle to the left side of the figure. The magnitude of $\vec{F}^E_-$ is greater than the magnitude of $\vec{F}^E_+$ because the negative end of the dipole is closer to the positively charged particle. Thus the vector sum of the forces exerted on the two ends is nonzero, and so the dipole experiences an acceleration whose magnitude and direction depend on its orientation with respect to the electric field. In addition, the forces create a torque about the dipole's center of mass. As in a uniform electric field:

> **A permanent electric dipole placed in an electric field is subject to a torque that tends to align the dipole moment with the direction of the electric field. If the field is uniform, the dipole's center of mass has zero acceleration; if the electric field is nonuniform, the dipole's center of mass has a nonzero acceleration.**



23.11 (*a*) Draw a free-body diagram for the dipole in Figure 23.18*a* and determine the direction of the dipole's center-of-mass acceleration. (*b*) Draw a free-body diagram for the dipole in Figure 23.18*b* and qualitatively describe the dipole's motion.

Figure 23.18 Extended free-body diagrams for permanent dipoles in nonuniform electric fields. The electric field shown is due to a positively charged particle to the left of the figure.

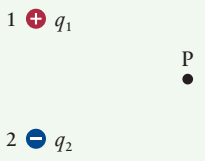


CONCEPTS

Self-quiz

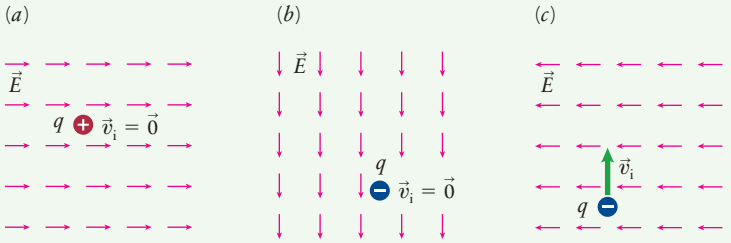
1. Suppose someone discovers that blue and yellow objects attract each other, that two blue objects repel each other, and that two yellow objects repel each other. The strength of this "chromatic interaction" is found to depend on color depth: The deeper the color, the greater the magnitude of the interaction. How would you define the magnitude and direction of the "chromatic field" of an object?
2. (*a*) Does an electrically neutral particle that has mass interact with an electric field? (*b*) Does a charged particle interact with a gravitational field?
3. The two particles in **Figure 23.19** have the same mass, carry charges of the same magnitude ($q_1 = -q_2 > 0$), and are equidistant from point P. (*a*) What is the electric field direction at P? (*b*) At P, what is the direction of the gravitational field due to the two particles? Ignore Earth's gravitational field.

Figure 23.19



4. Can electric and gravitational fields exist in the same place at the same time?
5. What are the directions of the acceleration of each particle in **Figure 23.20**? Describe the resulting motions.

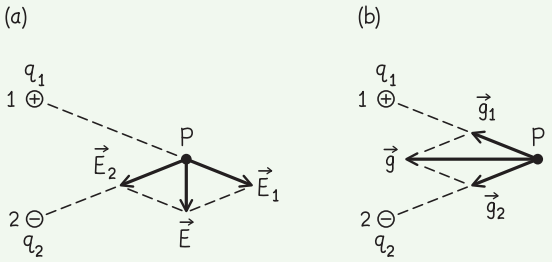
Figure 23.20



Answers

1. The gravitational field is defined as the gravitational force per unit of mass, with the field direction the same as the direction of the force. The electric field is defined as the electric force per unit of charge, with the field direction parallel to that of the force exerted on a positively charged particle. Therefore, the chromatic field can be defined as the chromatic force per unit of color, with the field direction parallel to that of the force exerted on a particle carrying some chosen color.
2. (*a*) No. Uncharged particles don't interact with electric fields. (Remember that a particle has no extent and therefore cannot be polarized.) (*b*) Yes, because any particle or object, charged or uncharged, interacts with a gravitational field.
3. (*a*) The electric field of particle 1 points away from the particle, which means that at P it points to the right and down. The electric field of particle 2 points toward the particle, meaning to the left and down at P. The vector sum of the electric fields at P therefore points straight down (**Figure 23.21a**). (*b*) The gravitational fields of the two particles point toward them from P, and so their vector sum points to the left (Figure 23.21*b*).

Figure 23.21



4. Yes. Consider a charged object near Earth's surface. This object is surrounded by an electric field, but it is also surrounded by Earth's (and to a lesser extent its own) gravitational field. These electric and gravitational fields exist in the same place at the same time.
5. (*a*) Recall from the discussion following Checkpoint 23.4 that the direction of the electric field at a given location is the same as the direction of the electric force exerted on a positively charged particle at that location. In Figure 23.20*a*, therefore, the positively charged particle experiences a force directed to the right. Because its initial velocity is zero, the particle moves in a straight line in the direction of the electric field. (*b*) The negatively charged particle in Figure 23.20*b* experiences an acceleration up the page, opposite the direction of the electric field. This particle moves in a straight line up the page. (*c*) The negatively charged particle in Figure 23.20*c* experiences an acceleration to the right, in the direction opposite the direction of the electric field. Because its initial velocity is perpendicular to the direction of the electric field, the particle travels in a parabolic trajectory up the page and curving to the right.

23.5 Electric field of a charged particle

In Section 23.2 we defined the **electric field** at a certain point P in space as the electric force experienced at P by a test particle carrying a charge q_t divided by the charge of the test particle:

$$\vec{E} \equiv \frac{\vec{F}_t^E}{q_t}. \tag{23.1}$$

The SI unit of electric field is the newton per coulomb (N/C).

Equation 23.1 requires no knowledge of the charge distribution that causes the electric field: It gives a prescription for determining the electric field at a given position in space. We can use Coulomb's law, however, to derive an expression for the electric field created at some point P due to a source particle carrying a charge q_s at position $\vec{r}_s$ (**Figure 23.22**). If we place a test particle carrying a charge q_t at P, Coulomb's law (Eq. 22.7) tells us that the force exerted on the test particle is

$$\vec{F}_{st}^E = k\frac{q_s q_t}{r_{st}^2}\hat{r}_{st}, \tag{23.2}$$

where $k = 9.0 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2$ (Eq. 22.5) is the proportionality constant that appears in Coulomb's law, r_{st} is the distance between the two particles, and $\hat{r}_{st}$ is a unit vector pointing from the source particle to the test particle. If we divide the electric force exerted by the source particle on the test particle by the charge q_t on the test particle, we obtain an expression for the electric field created by the source particle at P:

$$\vec{E}_s = \frac{\vec{F}_{st}^E}{q_t} = k\frac{q_s}{r_{st}^2}\hat{r}_{st}. \tag{23.3}$$

Because the test particle has nothing to do with this electric field, we can omit any reference to it by writing $\vec{r}_{st} = \vec{r}_{sP}$ and referring only to the position of point P:

$$\vec{E}_s(\text{P}) = k\frac{q_s}{r_{sP}^2}\hat{r}_{sP}. \tag{23.4}$$

This expression represents the electric field at P due to a source particle carrying a charge q_s at position $\vec{r}_s$.

As expected, the magnitude of the electric field at P is proportional to q_s, is independent of q_t, and decreases as the inverse square of the distance r_{sP} from the source particle. The direction of the electric field is outward (that is to say, in the direction given by $\hat{r}_{sP}$) when q_s is positive and inward (antiparallel to $\hat{r}_{sP}$) when q_s is negative.

Using the superposition principle, we now can determine the electric field due to a system of particles 1, 2, . . . carrying charges q_1, q_2, ... The combined electric field is the vector sum of the individual electric fields:

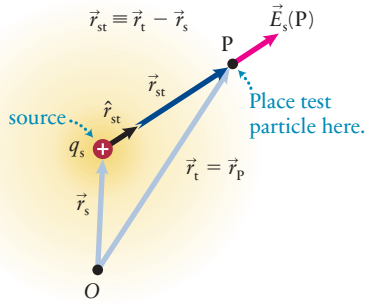
$$\vec{E} = \vec{E}_1 + \vec{E}_2 + \cdots = \Sigma k\frac{q_i\hat{r}_{iP}}{r_{iP}^2}. \tag{23.5}$$

Once the electric field at a certain position is known, the force exerted on any particle carrying charge q placed at that position can be found from

$$\vec{F}_p^E = q\vec{E}. \tag{23.6}$$

(Remember that we omit the "by" subscript on the force when the force is exerted by a field.) If q is positive, the force exerted on the particle is in the same direction as the electric field; if q is negative, the force exerted on the particle is in the direction opposite the direction of the electric field.

Figure 23.22 To determine the electric field at P generated by a charged source particle, we place a test particle at P.

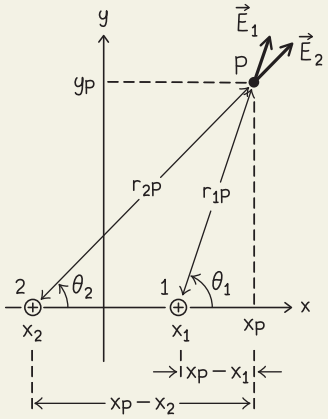


EXAMPLE 23.3 Electric field due to two charged particles

A point P is located at $x_P = 2.0$ m, $y_P = 3.0$ m. What are the magnitude and direction of the electric field at P due to a particle 1 carrying charge $q_1 = +10\ \mu$C and located at $x_1 = 1.0$ m, $y_1 = 0$ and a particle 2 carrying charge $q_2 = +20\ \mu$C and located at $x_2 = -1.0$ m, $y_2 = 0$?

❶ **GETTING STARTED** I begin by making a sketch of the situation (**Figure 23.23**). Each particle carries a positive charge, so the electric field due to each particle points away from the particle.

Figure 23.23



❷ **DEVISE PLAN** To determine the electric field $\vec{E}_P$ at P, I must take the vector sum of $\vec{E}_1$ and $\vec{E}_2$ at P. I can use Eq. 23.4 to calculate the magnitudes E_1 and E_2. To obtain the vector sum of the two fields, I add their x and y components.

❸ **EXECUTE PLAN** The distances from the particles to P are $r_{1P} = \sqrt{(x_P - x_1)^2 + y_P^2} = \sqrt{10\ \text{m}^2} = 3.2$ m and $r_{2P} = \sqrt{(x_P - x_2)^2 + y_P^2} = \sqrt{18\ \text{m}^2} = 4.2$ m. The magnitudes of the electric fields created by the particles at P are thus

$$E_1 = k\frac{|q_1|}{r_{1P}^2} = (9.0 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)\frac{(1.0 \times 10^{-5}\ \text{C})}{10\ \text{m}^2}$$
$$= 0.90 \times 10^4\ \text{N/C}$$

$$E_2 = k\frac{|q_2|}{r_{2P}^2} = (9.0 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)\frac{(2.0 \times 10^{-5}\ \text{C})}{18\ \text{m}^2}$$
$$= 1.0 \times 10^4\ \text{N/C}.$$

To calculate $\vec{E}_P$, I take the vector sum of $\vec{E}_1$ and $\vec{E}_2$ at P. In component form, I have

$$E_{Px} = E_{1x} + E_{2x} = E_1\cos\theta_1 + E_2\cos\theta_2$$
$$= E_1\frac{(x_P - x_1)}{r_{1P}} + E_2\frac{(x_P - x_2)}{r_{2P}}$$

$$E_{Py} = E_{1y} + E_{2y} = E_1\sin\theta_1 + E_2\sin\theta_2$$
$$= E_1\frac{y_P}{r_{1P}} + E_2\frac{y_P}{r_{2P}}.$$

Substituting the values given, I have

$$E_{Px} = (0.9 \times 10^4\ \text{N/C})\frac{1.0\ \text{m}}{3.2\ \text{m}} + (1.0 \times 10^4\ \text{N/C})\frac{3.0\ \text{m}}{4.2\ \text{m}}$$
$$= +1.0 \times 10^4\ \text{N/C}$$

$$E_{Py} = (0.9 \times 10^4\ \text{N/C})\frac{3.0\ \text{m}}{3.2\ \text{m}} + (1.0 \times 10^4\ \text{N/C})\frac{3.0\ \text{m}}{4.2\ \text{m}}$$
$$= +1.6 \times 10^4\ \text{N/C}.$$

Finally, I write this in vector form as

$$\vec{E}_P = (+1.0 \times 10^4\ \text{N/C})\hat{\imath} + (+1.6 \times 10^4\ \text{N/C})\hat{\jmath}.$$ ✔

❹ **EVALUATE RESULT** Both E_{Px} and E_{Py} are positive, as I expect based on my sketch. The magnitudes of $\vec{E}_1$ and $\vec{E}_2$ are comparable, which is what I would expect: Particle 2 carries twice the charge of particle 1, but the square of its distance to P is greater by a factor of 1.8.

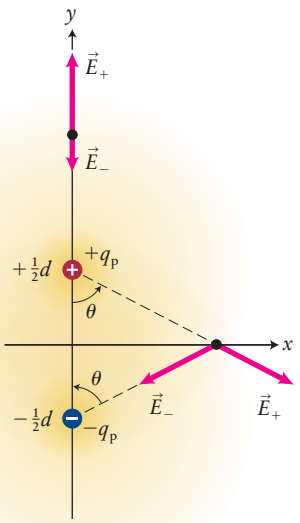


23.12 What is the magnitude of the electric force exerted by the electric field on an electron placed at point P in Figure 23.23? (Assume that the only interaction is electrostatic.) What is the initial acceleration of the electron if it is released from rest from that point? [$e = 1.6 \times 10^{-19}$ C; $m_e = 9.1 \times 10^{-31}$ kg.]

23.6 Dipole field

Next we examine the electric field due to a permanent electric dipole. **Figure 23.24** on the next page shows a dipole that consists of a particle carrying a charge $+q_p$ at $x = 0$, $y = +\frac{1}{2}d$, and another particle carrying a charge $-q_p$ at $x = 0$, $y = -\frac{1}{2}d$, where d is the distance between the two particles. The charge q_p of the positively charged pole is called the *dipole charge*, and the distance d is called the *dipole separation*. Each particle creates an electric field at all positions in space, so the two fields overlap

Figure 23.24 Calculating the electric field due to a dipole.



everywhere. We can determine the combined electric field at any position by adding the two fields vectorially. Let us do this for two general locations: anywhere along the x axis and anywhere along the y axis.

Along the x axis, which bisects the dipole, the magnitudes of the electric fields due to the two ends of the dipole are equal:

$$E_+ = E_- = k\frac{q_p}{x^2 + (d/2)^2}. \tag{23.7}$$

The x components of these two electric fields point in opposite directions and so add to zero. The magnitude of the combined electric field is thus equal to the sum of the y components:

$$E_y = E_{+y} + E_{-y} = -(E_+ + E_-)\cos\theta$$

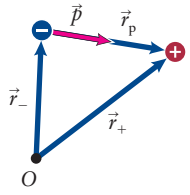
$$= -\left(2k\frac{q_p}{x^2+(d/2)^2}\right)\left(\frac{d/2}{[x^2+(d/2)^2]^{1/2}}\right) = -k\frac{q_p d}{[x^2+(d/2)^2]^{3/2}}. \tag{23.8}$$

The product $q_p d$ is a measure of the strength of the dipole and is the magnitude of the dipole moment introduced in Section 23.4. To specify both the strength and the orientation of the dipole we can write this quantity as a vector, called the **dipole moment:**

$$\vec{p} \equiv q_p \vec{r}_p, \tag{23.9}$$

where $\vec{r}_p \equiv \vec{r}_{-+} = \vec{r}_+ - \vec{r}_-$ is the position of the positively charged particle relative to the negatively charged particle (and so $d = |\vec{r}_p|$). Because q_p is always taken to be positive, the dipole moment $\vec{p}$ points in the same direction as $\vec{r}_p$: along the axis of the dipole (the line that passes through the center of each particle), in the direction from the negative to the positive pole (**Figure 23.25**). Large permanent dipole moments can be caused either by a large dipole separation d or by a large dipole charge q_p. Conceptually you can think of the magnitude of the dipole moment as a measure of how strongly the dipole wants to align itself in the direction of an electric field. The SI unit of dipole moment is C · m.

Figure 23.25 The dipole moment $\vec{p}$ points along the axis of the dipole from the negative to the positive pole.



For distances far from the dipole ($x \gg d/2$), we may ignore $d/2$, and so $[x^2 + (d/2)^2]^{3/2} \to x^3$. Equation 23.8 thus becomes

$$E_y \approx -k\frac{p}{|x^3|} \quad \text{(far from dipole along the positive } x \text{ axis)}. \tag{23.10}$$

The right side of this equation is negative for both positive and negative x, and so anywhere along the x axis the dipole's electric field $\vec{E}$ points in the negative y direction, opposite the direction of the dipole moment. Equation 23.10 also shows that the magnitude of the electric field is inversely proportional to x^3, in contrast to the electric field of a charged particle, which is inversely proportional to x^2 (Eq. 23.4). The reason the electric field of a dipole approaches zero faster as x increases is that the angle between $\vec{E}_+$ and $\vec{E}_-$ in Figure 23.24 approaches 180° as x increases, and so the electric fields of the two poles tend to cancel each other more and more.

Along the y axis, the electric field created by either end of the dipole is directed along the y axis. Thus to determine the y component of the electric field of the dipole at any position along the y axis, we must add the y components of the fields from each particle. For $y > +d/2$:

$$E_y = E_{+y} + E_{-y} = k\frac{q_p}{[y-(d/2)]^2} - k\frac{q_p}{[y+(d/2)]^2}. \tag{23.11}$$

QUANTITATIVE TOOLS

After some algebra, this can be rewritten in the form

$$E_y = k\frac{q_p}{y^2}\left[\left(1-\frac{d}{2y}\right)^{-2}-\left(1+\frac{d}{2y}\right)^{-2}\right] \qquad (y>+d/2). \quad (23.12)$$

For distances far from the dipole, $y \gg +d/2$, so we can use the binomial series expansion, which states that for $x \ll 1$, $(1+x)^n \approx 1+nx$ (see Appendix B). Applying this expansion to the two terms inside the square brackets in Eq. 23.12, we get

$$E_y \approx k\frac{q_p}{y^2}\left[\left(1+2\frac{d}{2y}\right)-\left(1-2\frac{d}{2y}\right)\right]$$

$$= k\frac{q_p}{y^2}\left[\frac{2d}{y}\right] = 2k\frac{q_p d}{y^3} = 2k\frac{p}{y^3} \qquad (y \gg d/2). \quad (23.13)$$

The right side of this equation has the same algebraic sign as y, so the dipole's electric field $\vec{E}$ points in the positive y direction. (Carrying out the same calculation for $y<-d/2$, you can show that the electric field still points in the positive y direction. In between the two charged particles, the electric field points in the negative y direction.) The magnitude of the electric field is inversely proportional to y^3—just as along the x axis, the electric fields of each of the two poles tend to cancel each other more and more as the distance from the dipole increases. One can show that the electric field of the dipole depends on $1/r^3$ for all positions far from the dipole (where r is the distance between the point under consideration and the center of the dipole). The reason is that the electric fields of the positive and negative ends of the dipole partially cancel each other, and this cancellation becomes more complete far from the dipole: The farther you are from the dipole, the smaller the separation between the charged particles appears to be.

23.13 The magnitude of the electric field created by dipole A at a certain point P is E_A. If the dipole is replaced with another dipole B that has its dipole moment oriented in the same direction, the magnitude of the electric field at point P is found to be greater: $E_B > E_A$. Which dipole has the greater dipole moment? For which of these two dipoles is the dipole charge q_p greater?

23.7 Electric fields of continuous charge distributions

So far we have dealt with only charged particles because Coulomb's law applies only to charged particles. However, most charged objects of interest—from charged combs to electrical components—are not particles. Instead, they are extended bodies. Although every macroscopic object consists of very large numbers of charged particles—protons and electrons—it is not practical to calculate the individual field of each of these particles and then add them vectorially. Instead, we shall treat any macroscopic charged object as having a continuous charge distribution and calculate the electric field created by the object by dividing the charge distribution on the object into infinitesimally small segments that may be considered charged source particles carrying a charge dq_s. For the charged macroscopic object shown in Figure 23.26, for example, we can use Coulomb's law to obtain the infinitesimal portion of the electric field at point P contributed by a segment:

$$d\vec{E}_s(\mathrm{P}) = k\frac{dq_s}{r_{sP}^2}\hat{r}_{sP}. \qquad (23.14)$$

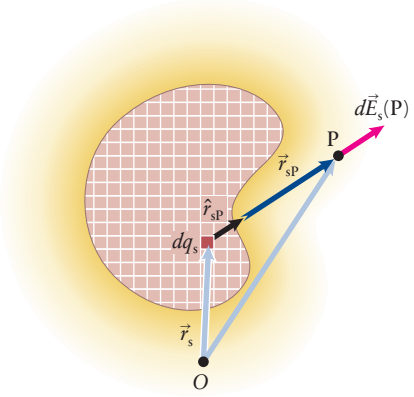


Figure 23.26 To calculate the electric field created at P by a continuous charge distribution, we divide the distribution into infinitesimally small segments that can be treated as charged source particles carrying charge dq_s.

Using the principle of superposition, we can then sum the contributions of all the segments that make up the object. Because the segments are infinitesimally small, this sum corresponds to an integral:

$$\vec{E} = \int d\vec{E}_s = k\int \frac{dq_s}{r_{sP}^2}\hat{r}_{sP}. \tag{23.15}$$

In order to evaluate this integral, we must express dq_s, $1/r_{sP}^2$, and $\hat{r}_{sP}$ in terms of the same coordinate(s). To do so, it is necessary to express the charge on the object in terms of a **charge density**—the amount of charge per unit of length, per unit of surface area, or per unit of volume. For a one-dimensional object, such as a thin charged wire of length ℓ carrying a charge q uniformly distributed along the wire, the *linear charge density*—the amount of charge per unit of length (in coulombs per meter)—is given by

$$\lambda \equiv \frac{q}{\ell} \quad \text{(uniform charge distribution)}. \tag{23.16}$$

For uniformly charged two-dimensional objects, we use the *surface charge density*—the amount of charge per unit of area (in coulombs per square meter). For example, the surface charge density of a flat plate of area A carrying a uniformly distributed charge q is

$$\sigma \equiv \frac{q}{A} \quad \text{(uniform charge distribution)}. \tag{23.17}$$

For a uniformly charged three-dimensional object, we use the *volume charge density:*

$$\rho \equiv \frac{q}{V} \quad \text{(uniform charge distribution)}, \tag{23.18}$$

which gives the amount of charge per cubic meter.

The Procedure box below provides some helpful steps for carrying out the integral in Eq. 23.15, and the next four examples show how to put the procedure into practice.

PROCEDURE: Calculating the electric field of continuous charge distributions by integration

To calculate the electric field of a continuous charge distribution, you need to evaluate the integral in Eq. 23.15. The following steps will help you evaluate the integral.

1. Begin by making a sketch of the charge distribution. Mentally divide the distribution into small segments. Indicate one such segment that carries a charge dq_s in your drawing.
2. Choose a coordinate system that allows you to express the position of the segment in terms of a minimum number of coordinates (x, y, z, r, or θ). These coordinates are the integration variables. For example, use a radial coordinate system for a charge distribution with radial symmetry. Unless the problem specifies otherwise, let the origin be at the center of the object.
3. Draw a vector showing the electric field caused by the segment at the point of interest. Examine how the components of this vector change as you vary the position of the segment along the charge distribution. Some components may cancel, which greatly simplifies the calculation. If you can determine the direction of the resulting electric field, you may need to calculate only one component. Otherwise express $\hat{r}_{sP}$ in terms of your integration variable(s) and evaluate the integrals for each component of the field separately.
4. Determine whether the charge distribution is one-dimensional (a straight or curved wire), two-dimensional (a flat or curved surface), or three-dimensional (any bulk object). Express dq_s in terms of the corresponding charge density of the object and the integration variable(s).
5. Express the factor $1/r_{sP}^2$, where r_{sP} is the distance between dq_s and the point of interest, in terms of the integration variable(s).

At this point you can substitute your expressions for dq_s and $1/r_{sP}^2$ into Eq. 23.15 and carry out the integral (or component integrals), using what you determined about the direction of the electric field (or substituting your expression for $\hat{r}_{sP}$).

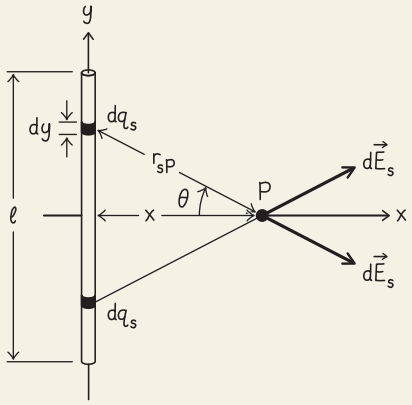
QUANTITATIVE TOOLS

EXAMPLE 23.4 Electric field created by a uniformly charged thin rod

A thin rod of length ℓ carries a uniformly distributed charge q. What is the electric field at a point P along a line that is perpendicular to the long axis of the rod and passes through the rod's midpoint?

❶ **GETTING STARTED** I begin by making a sketch of the situation. After drawing a set of axes, I place the rod along the y axis, with the origin at the rod center and point P on the positive x axis (**Figure 23.27**).

Figure 23.27



❷ **DEVISE PLAN** The word *thin* implies that I can treat the rod as a one-dimensional object. Because the rod is uniformly charged, I can thus use Eq. 23.16 to determine the linear charge density along the rod. To determine the electric field at P, I divide the rod lengthwise into a large number of infinitesimally small segments, each of length dy. Each segment contributes to the electric field at P an amount given by Eq. 23.14. For each segment above the x axis there is a corresponding segment below the axis at the same distance from P. The y components of the electric fields $d\vec{E}_s$ due to these two segments add up to zero, so I need to calculate only the x component dE_{sx}. To get the electric field created by the entire rod, I use Eq. 23.15 to integrate my result over the length of the rod.

❸ **EXECUTE PLAN** The charge dq_s on each segment dy is $dq_s = \lambda\, dy = (q/\ell)\, dy$. The x component of the electric field created by each segment at P is thus

$$dE_{sx} = k\frac{dq_s}{r_{sP}^2}\cos\theta = k\frac{q}{\ell r_{sP}^2}\cos\theta\, dy, \tag{1}$$

where θ is the angle between the x axis and the line that connects the segment dy with P. Both θ and r_{sP} depend on the position y of the segment, so I must choose one integration variable and express the others in terms of that variable. I choose θ as the integration variable, which means I must express the factor dy/r_{sP}^2 in Eq. 1 in terms of θ. Using trigonometry, I have

$$\cos\theta = \frac{x}{r_{sP}} \tag{2}$$

and

$$\tan\theta = \frac{y}{x}, \tag{3}$$

where x is the x coordinate of point P. Differentiating Eq. 3 yields

$$dy = x\, d(\tan\theta) = \frac{x}{\cos^2\theta}d\theta. \tag{4}$$

Next I divide Eq. 4 by r_{sP}^2 to obtain the factor dy/r_{sP}^2 I need. I use r_{sP}^2 on the left, but on the right I use Eq. 2 to write r_{sP}^2 in the form $x^2/\cos^2\theta$, yielding

$$\frac{dy}{r_{sP}^2} = \left(\frac{\cos^2\theta}{x^2}\right)\left(\frac{x}{\cos^2\theta}d\theta\right) = \frac{1}{x}d\theta.$$

Substituting this result into Eq. 1 and integrating over the entire rod yield

$$E_x = k\frac{q}{\ell}\int_{-\theta_{max}}^{+\theta_{max}}\frac{\cos\theta}{x}d\theta = \frac{kq}{\ell x}\int_{-\theta_{max}}^{+\theta_{max}}\cos\theta\, d\theta$$
$$= \frac{kq}{\ell x}\sin\theta\Big|_{-\theta_{max}}^{+\theta_{max}} = \frac{2kq}{\ell x}\sin\theta_{max},$$

where θ_{max}, the maximum value of θ, is the angle between the x axis and the line that connects the top end of the rod with P. Substituting $\sin\theta_{max} = y/r_{sP} = \frac{1}{2}\ell/\sqrt{(\ell/2)^2 + x^2}$ finally yields

$$E_x = \frac{kq}{x\sqrt{\ell^2/4 + x^2}};\ E_y = 0;\ E_z = 0. \checkmark$$

❹ **EVALUATE RESULT** Very far from the rod along the positive x axis, $x \ll \ell$, so the rod looks like a particle. In this case, I can ignore the ℓ^2 term in the denominator and my result becomes identical to Eq. 23.4, the equation for a particle that carries a charge q ($E = kq/r^2$, or using the symbols from this problem, $E_x = kq/x^2$).

When P is very close to the rod, $x \ll \ell$, I can ignore the x^2 term in the denominator and write

$$E_x = \frac{2k(q/\ell)}{x} = \frac{2k\lambda}{x}. \tag{5}$$

In this $x \ll \ell$ case, the distance from P to either end of the rod is much greater than the distance from P to the closest point on the rod (which is the rod midpoint, located at the origin), and thus the rod essentially looks "infinitely long" to an observer at P. Indeed, Eq. 5 shows that my result no longer depends on ℓ.

I also note that the rod's electric field is now inversely proportional to x rather than x^2. I saw in Chapter 17 that the amplitudes of waves that spread out in three dimensions are inversely proportional to x^2, whereas the amplitudes of waves that spread out in two dimensions are inversely proportional to x. My result therefore make sense because the electric field that emanates from a charged particle "spreads out" in three dimensions, but the field that emanates from an infinitely long charged rod spreads out in just two dimensions.

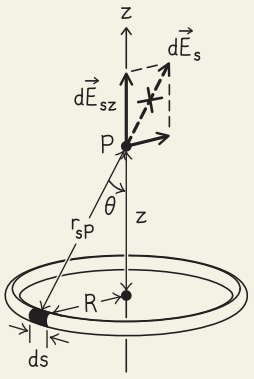
QUANTITATIVE TOOLS

EXAMPLE 23.5 Electric field created by a uniformly charged thin ring

A thin ring of radius R carries a uniformly distributed charge q. What is the electric field at point P along an axis that is perpendicular to the plane of the ring and passes through its center?

1 GETTING STARTED I begin by making a sketch of the situation. I let the ring be in the xy plane, with the origin at the center of the ring, and I place P on the positive z axis (**Figure 23.28**).

Figure 23.28



2 DEVISE PLAN Because the ring is thin, I can use Eq. 23.16 to determine its linear charge density. To determine the electric field at P, I divide the ring into a large number of infinitesimally small segments, each of arc length ds. Each segment contributes to the electric field an amount given by Eq. 23.14. Because all segments are at the same distance from P, all contribute an electric field $d\vec{E}_s$ of the same magnitude. As shown in my sketch, the contribution $d\vec{E}_s$ makes an angle θ with the z axis, and so each segment ds produces a component parallel to the z axis and a component perpendicular to it. For each pair of segments on opposite sides of the ring, the components of $d\vec{E}_s$ perpendicular to the z axis add up to zero, and so I am concerned only with the z components.

To get the electric field created by the ring, I can use Eq. 23.15 to integrate my $d\vec{E}_s$ result over the circumference of the ring.

3 EXECUTE PLAN Each segment carries a charge $dq_s = \lambda\, ds$, where $\lambda = q/2\pi R$ is the linear charge density along the ring. The magnitude of each segment's contribution to the electric field is, from Eq. 23.14,

$$dE_s = k\frac{dq_s}{r_{sP}^2} = k\frac{\lambda\, ds}{r_{sP}^2} = k\frac{\lambda\, ds}{z^2 + R^2}.$$

For the z component of $d\vec{E}_s$, I see from Figure 23.28 that the angle between the vectors $d\vec{E}_{sz}$ and $d\vec{E}_s$ is also θ, and so $\cos\theta = dE_{sz}/dE_s$. Then, combining this expression with the relationship

$$\cos\theta = \frac{z}{r_{sP}} = \frac{z}{\sqrt{z^2 + R^2}},$$

I have

$$dE_{sz} = \cos\theta\, dE_s = \left[\frac{z}{\sqrt{z^2 + R^2}}\right]\left[k\frac{\lambda\, ds}{z^2 + R^2}\right] = k\frac{z\lambda}{[z^2 + R^2]^{3/2}}ds.$$

To determine the electric field created by the ring, I must integrate the contributions around the ring, from $s = 0$ to $s = 2\pi R$. Because k, z, R, and λ are all independent of s, I can move everything out of the integral except ds:

$$E_z = \int dE_{sz} = k\frac{z\lambda}{[z^2 + R^2]^{3/2}}\int_0^{2\pi R} ds = k\frac{z\lambda(2\pi R)}{[z^2 + R^2]^{3/2}}.$$

Because $\lambda = q/2\pi R$, the term $\lambda(2\pi R)$ is equal to the charge q on the ring, so I get for the z component of the electric field along the axis perpendicular to the plane of the ring and passing through the ring center

$$E_x = 0;\ E_y = 0;\ E_z = k\frac{qz}{[z^2 + R^2]^{3/2}}. \checkmark$$

4 EVALUATE RESULT Very far from the ring my expression for E_z should become the same as that for a charged particle. Indeed, when $z \gg R$, I can ignore the R^2 term in my result, and so

$$E_z \approx k\frac{qz}{[z^2]^{3/2}} = k\frac{qz}{z^3} = k\frac{q}{z^2},$$

as I expect.

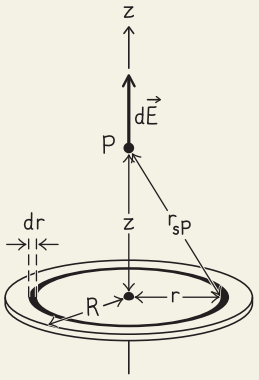
At the center of the ring, $z = 0$ and so my expression yields $E_z = 0$. This result is reasonable because at the center of the ring the electric forces exerted by segments on opposite sides of the ring on a charged test particle add to zero. When the vector sum of these forces is zero, the electric field must be zero also.

EXAMPLE 23.6 Electric field created by a uniformly charged disk

A thin disk of radius R carries a uniformly distributed charge. The surface charge density on the disk is σ. What is the electric field at a point P along the perpendicular axis through the disk center?

1 GETTING STARTED I begin with a sketch, placing the disk in the xy plane, with the disk center at the origin. I let point P lie on the positive z axis (**Figure 23.29**).

Figure 23.29



2 DEVISE PLAN Because of the circular symmetry of the disk, I divide it into a large number of ring-shaped segments, each of radius r and width dr. The charge on each ring is the product of the ring surface area (circumference times width) and the surface charge density: $dq_s = (2\pi r)dr\,\sigma$. The

contribution of each ring to the electric field at P is given by the expression for E_z I obtained for a uniformly charged thin ring in Example 23.5 with $2\pi r\,\sigma\,dr$ substituted for q and r substituted for R. So all I need to do is integrate over the entire disk.

❸ **EXECUTE PLAN** Substituting $q = 2\pi r\,\sigma\,dr$ and $R = r$ into the Example 23.5 expression for E_z and integrating the result over the disk from $r = 0$ to $r = R$, I have

$$E_z = \int dE_z = k\int_0^R \frac{2\pi r\,\sigma z}{(z^2+r^2)^{3/2}}dr.$$

Because σ and z are independent of r, I can move them out of the integral:

$$E_x = 0;\ E_y = 0;$$

$$E_z = k\pi\sigma z\int_0^R \frac{2r\,dr}{(z^2+r^2)^{3/2}}$$

$$= k\pi\sigma z\int_0^R \frac{d(r^2)}{(z^2+r^2)^{3/2}} = k\pi\sigma z\frac{-2}{(z^2+r^2)^{1/2}}\Bigg|_0^R$$

$$= 2k\pi\sigma z\left[\frac{1}{(z^2)^{1/2}} - \frac{1}{(z^2+R^2)^{1/2}}\right]. \checkmark \qquad (1)$$

❹ **EVALUATE RESULT** Let me evaluate my result for $z \gg R$, where, to an observer at P, the disk looks like a particle. For positive z, I can write the E_z in Eq. 1 as

$$E_z = 2k\pi\sigma\left[1 - \frac{z}{(z^2+R^2)^{1/2}}\right]. \qquad (2)$$

From the binomial series expansion (see Appendix B), I get in the case that $z \gg R$,

$$\frac{z}{(z^2+R^2)^{1/2}} = \left(1+\frac{R^2}{z^2}\right)^{-1/2} \approx 1 - \tfrac{1}{2}\frac{R^2}{z^2}.$$

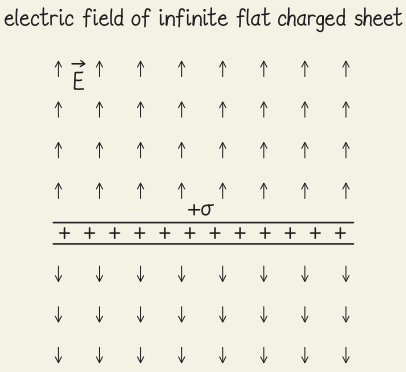
Substituting this result into Eq. 2 and writing $q = \sigma(\pi R^2)$ for the charge on the disk, I get

$$E_z = 2k\pi\sigma\left(\tfrac{1}{2}\frac{R^2}{z^2}\right) = k\frac{\sigma\pi R^2}{z^2} = k\frac{q}{z^2},$$

which is the result for a charged particle, as I expect.

I can also evaluate my result for $z \approx 0$, where, to an observer at that location, the disk looks like it has an infinite radius (that is to say, it looks like an infinite flat sheet). In that case, the second term inside the brackets in Eq. 2 is zero and $E_z = 2k\pi\sigma$. This tells me that the electric field of an infinite flat charged sheet is independent of z and constant throughout space, as I have sketched in **Figure 23.30**. (In other words, it is uniform.) While this lack of dependence on z is somewhat counterintuitive, it agrees with what I concluded earlier: The electric field that emanates from a charged particle "spreads out" in three dimensions and its amplitude is inversely proportional to the square of the distance from the particle, whereas the electric field that emanates from an infinitely long charged rod spreads out in just two dimensions and its amplitude is inversely proportional to the distance from the rod. As my sketch shows, the electric field that emanates from an infinite plane can't spread out at all (if the plane is truly infinite), and therefore its amplitude is independent of distance.

Figure 23.30



23.14 (a) Describe the electric field between two infinitely large parallel charged sheets if the charge density of one sheet is $+\sigma$ and that of the other is $-\sigma$. (*b*) Describe the electric field outside the sheets.

EXAMPLE 23.7 Electric field created by a uniformly charged sphere

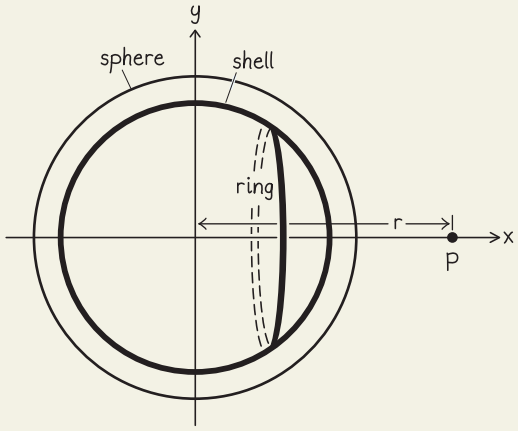
A solid sphere of radius R carries a fixed, uniformly distributed charge q. Exploiting the analogy between Newton's law of gravity and Coulomb's law, use the result obtained in Section 13.8 to obtain an expression for the magnitude of the electric field created by the sphere at a point P outside the sphere.

❶ **GETTING STARTED** To determine the electric field magnitude at point P, I need to determine the magnitude of the electric force exerted by the sphere on a test particle carrying a charge q_t at P and then divide that force magnitude by q_t.

❷ **DEVISE PLAN** I can follow the same procedure as in Section 13.8 to calculate the gravitational force of a spherical object: I first divide the sphere into a series of thin concentric shells that resemble the layers in an onion and then

(Continued)

Figure 23.31



divide each shell into a series of vertical rings (**Figure 23.31**). I then calculate the contribution of each ring to the electric field at P and integrate first over each shell and then over the sphere. The expression I get for F^E_{sphere} must be of the same form as that for the gravitational sphere, F^G_{sphere} (Eq. 13.37), because the gravitational force and the electric force are both inversely proportional to the square of the distance between the interacting particles. So all I need to do is replace G in Eq. 13.37 by k, M_{sphere} by q, and m by q_t to obtain the magnitude of the electric force exerted by the sphere on the test particle. To obtain an expression for the electric field, I then divide the result by q_t.

❸ **EXECUTE PLAN** In analogy to Eq. 13.37, I write

$$F^G_{\text{sp}} = G\frac{mM_{\text{sphere}}}{r^2} \rightarrow F^E_{\text{st}} = k\frac{q_t q}{r^2},$$

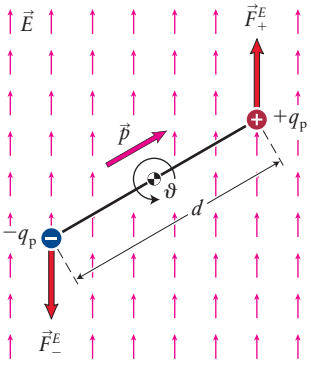
where r is the distance from the center of the sphere to P. To obtain the magnitude of the electric field, I divide the electric force by q_t:

$$E_{\text{sphere}} = k\frac{q}{r^2}. \checkmark$$

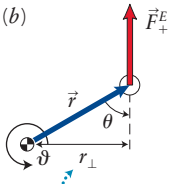
❹ **EVALUATE RESULT** Comparing my result with Eq. 23.3 (in scalar form), $E_s = kq_s/r^2_{st}$, shows that outside a uniformly charged solid sphere the magnitude of the electric field is the same as that surrounding a particle carrying the same charge and located at the center of the sphere. The result is independent of the radius R of the sphere and similar to the result obtained in Section 13.8: A solid sphere exerts a gravitational force as if the entire mass of the sphere were concentrated at the center. It makes sense that I obtain a similar result for the charged sphere because the gravitational force and the electric force are both proportional to $1/r^2$.

Figure 23.32 The torque on an electric dipole caused by an electric field tends to align the dipole moment $\vec{p}$ with the direction of the electric field.

(*a*) Electric dipole in electric field

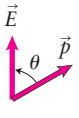


(*b*)



Lever arm $\vec{r}_\perp$ depends on angle θ of dipole with respect to electric field.

(*c*)



Torque on dipole is vector product $\Sigma\vec{\tau} = \vec{p} \times \vec{E}$.

23.15 How does the electric field inside a uniformly charged sphere vary with distance from the sphere center? [Hint: What is the electric field inside a hollow uniformly charged sphere?]

23.8 Dipoles in electric fields

Let us end this chapter by considering the forces exerted by electric fields on dipoles. **Figure 23.32*a*** shows a dipole consisting of two particles that carry charges of equal magnitude but opposite sign connected by a rod of length d; the dipole makes an angle θ with a uniform electric field $\vec{E}$ created by some unseen distant source. As we saw in Section 23.4, the forces exerted by the electric field on the charged ends of the dipole are equal in magnitude but opposite in direction, and so the vector sum of the forces exerted on the dipole is zero. Consequently the acceleration of the center of mass of the dipole is zero. Because the forces are exerted on opposite ends of the dipole, however, they create torques that cause the dipole to rotate counterclockwise about its center of mass. Figure 23.32*b* shows that the force exerted on the positive end causes a counterclockwise torque of magnitude

$$\tau_+ = r_\perp F^E_+ = \left(\tfrac{1}{2}d\sin\theta\right)(q_pE), \qquad (23.19)$$

where $r_\perp$ is the lever arm of the force. The force exerted on the negative end causes an identical torque because the lever arm and the magnitude of the force are the same. The electric field thus causes a torque on the dipole equal to

$$\Sigma\tau_\vartheta = 2\left(\tfrac{1}{2}d\sin\theta\right)(q_pE) = (q_pd)E\sin\theta \equiv pE\sin\theta. \qquad (23.20)$$

QUANTITATIVE TOOLS

This can be written in vectorial form as

$$\Sigma\vec{\tau} = \vec{p} \times \vec{E}, \tag{23.21}$$

where $\vec{p}$ is the dipole moment, which by definition points from the negative end of the dipole to the positive end and whose magnitude is given by Eq. 23.9. According to the right-hand rule (see Section 12.4), the vector product $\vec{p} \times \vec{E}$ in Eq. 23.21 gives a torque that points out of the plane of the drawing in Figure 23.32*c*. As we saw in Section 12.8, such a torque indeed causes a counterclockwise rotation. The torque on the dipole is maximum when the dipole moment is perpendicular to the electric field and zero when it is parallel or antiparallel to the electric field.

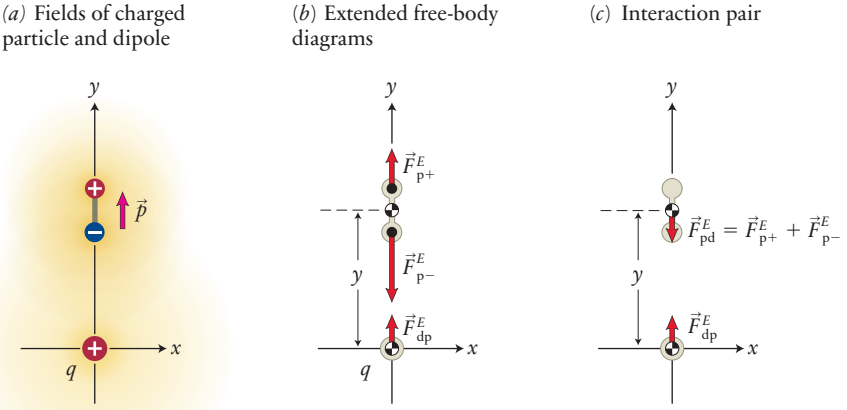
23.16 Is Eq. 23.20 valid if the center of mass is not halfway between the two charged ends of the dipole?

As we saw in Section 23.4, the vector sum of the forces exerted on dipoles in nonuniform electric fields is not zero. Consider, for example, the situation illustrated in **Figure 23.33*a***. A dipole with its dipole moment $\vec{p}$ aligned along the y axis is placed in the nonuniform electric field generated by a particle carrying a charge q and located at the origin. Because the distance between the negative end of the dipole and the particle is smaller than the distance between the positive end and the particle, the magnitude of the attractive force $\vec{F}^E_{p-}$ on the negative end is greater than the repulsive force $\vec{F}^E_{p+}$ on the positive end. Consequently the vector sum of the forces exerted by the nonuniform field on the dipole is nonzero, and the dipole is attracted to the particle. How does this attraction vary with the position y of the dipole?

To answer this question, we can write an expression for the vector sum of the forces $\Sigma\vec{F}^E_d$ exerted on the two ends of the dipole and examine how this sum varies with y. Alternatively, we can calculate the force $\vec{F}^E_{dp}$ exerted by the dipole on the particle, using our results from Section 23.6. This force and the vector sum of the forces exerted by the particle on the dipole form an interaction pair, and so their magnitudes are the same. Equation 23.13 tells us that, along the dipole axis, the magnitude of the electric field created by the dipole is $2k(p/y^3)$, and so the magnitude of the force exerted by the dipole on the particle is

$$F^E_{dp} = qE_d = 2k\frac{pq}{y^3}. \tag{23.22}$$

Figure 23.33 A dipole interacts with a charged particle.



QUANTITATIVE TOOLS

The magnitude of the force exerted by the particle on the dipole, being equal in magnitude to F^E_{dp}, is thus

$$F^E_{pd} = F^E_{p-} - F^E_{p+} = 2k\frac{pq}{y^3}. \tag{23.23}$$

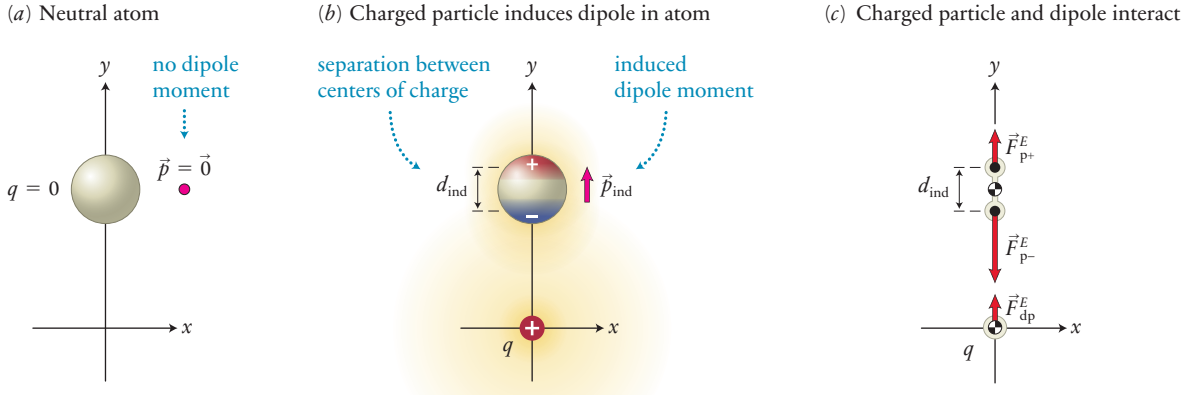
Like the electric field of a dipole, the forces between a charged object carrying charge q and a dipole is inversely proportional to the cube of the distance between them.

23.17 How does doubling each of the following quantities affect the force between a dipole and a particle placed near the dipole and carrying charge q: (*a*) the charge q, (*b*) the dipole separation d of the dipole, (*c*) the dipole charge q_p, (*d*) the distance between the dipole and the charged particle?

As we saw in Chapter 22, electrically neutral objects interact with a charged object because they become polarized in the presence of the charged object. Consider an isolated neutral atom. The centers of the atom's positive and negative charge distributions coincide, and therefore the atom's dipole moment is zero: $d = 0$ and so $\vec{p} = \vec{0}$ (**Figure 23.34*a***). The presence of an external electric field—that is, an electric field created by some other charged object—causes a separation between the positive and negative charge centers and so induces a dipole moment (Figure 23.34*b*). To understand the interaction between charged objects and neutral ones, we must therefore study the interaction between a charged particle and what is called an **induced dipole**. The first question to ask is: How does the magnitude of the induced dipole moment depend on the presence of a charged particle?

When a neutral atom is placed in an electric field $\vec{E}$, it is found that, as long as the electric forces exerted by that field on the charged particles in the atom are not too large, the induced dipole separation d_{ind} in the atom obeys Hooke's law. In other words, the induced dipole separation is proportional to the magnitude of the applied electric force, $F^E_d = cd_{ind}$, with c being the "spring constant" of the atom. We can rewrite this as $d_{ind} = (1/c)F^E_d$, and because d_{ind} is proportional to the magnitude of the induced dipole moment p_{ind} and the magnitude of the force exerted on the dipole is proportional to the magnitude E of the electric field at the

Figure 23.34 A charged particle induces a dipole in an electrically neutral atom.



position of the dipole, the **induced dipole moment** is proportional to the field at the position of the dipole:

$$\vec{p}_{\text{ind}} = \alpha\vec{E} \quad (E \text{ not too large}), \tag{23.24}$$

where α, the **polarizability** of the atom, is a constant that expresses how easily the charge distributions in the atom are displaced from each other. The SI unit of polarizability is $C^2 \cdot m/N$.

23.18 Given that the induced dipole moment $\vec{p}_{\text{ind}}$ points from the negative to the positive end of an induced dipole and the electric field $\vec{E}$ displaces the positive charge center in the direction of the electric field and the negative charge center in the opposite direction, do you expect the polarizability α to be positive or negative?

The electric field of a charged particle is given by Eq. 23.4, so the magnitude of the induced dipole moment is proportional to the inverse square of the distance between the particle and the dipole:

$$p_{\text{ind}} = \alpha E = \alpha\, k\frac{q}{y^2}. \tag{23.25}$$

In contrast, the dipole moment of a permanent dipole is constant.

We can now substitute the induced-dipole result of Eq. 23.25 into Eq. 23.23 to determine the force exerted by a charged particle on an induced dipole:

$$F^E_{\text{pd}} = 2k\frac{p_{\text{ind}}q}{y^3} = \alpha\frac{2k^2q^2}{y^5}. \tag{23.26}$$

This result shows that the interaction between a charged particle and a polarized object depends much more strongly on the distance between them $(1/y^5)$ than does the interaction between two charged objects $(1/y^2)$. You may have noticed this in Chapter 22 when comparing the attraction between two charged strips of tape with the attraction between a charged strip and a neutral object.* As the neutral object approaches the charged strip, the force varies so fast with distance that it is often difficult to prevent the tape from sticking to the neutral object.

23.19 (*a*) How does doubling the charge q_A carried by an object A affect the force exerted by A on another charged particle? (*b*) How does doubling q_A affect the force exerted by A on an induced dipole? (*c*) Explain why your answers to parts *a* and *b* are the same or different. (*d*) Can the force exerted by a charged particle cause a torque on an induced dipole?

*Try it! Pull two strips of transparent tape out of a dispenser, suspend one from the edge of a table and then move the other slowly toward it. Notice how the interaction between the strips varies relatively smoothly as a function of separation. Next, move your hand slowly toward the suspended strip and note how the force increases rapidly.

QUANTITATIVE TOOLS

Chapter Summary

Electric fields (Sections 23.1–23.3, 23.5)

Concepts In the field model of long-range interactions, interacting object A fills the space around itself with a **field** (or **interaction field**). When an object B is placed in this field, the field acts on B. If A and B have mass, the field is a *gravitational field;* if they are charge carriers, it is an *electric field*.

A **test particle** is one whose mass and/or charge is so small that the particle does not perturb the objects generating the field being measured.

The direction of the **electric field** at a given location is the same as the direction of the electric force exerted on a positive test particle at that location.

Quantitative Tools If a test particle located at position P and carrying charge q_t experiences a force $\vec{F}_t^E$, then the **electric field** $\vec{E}$ (N/C) at P is

$$\vec{E} \equiv \frac{\vec{F}_t^E}{q_t}. \tag{23.1}$$

If the source particle is at position $\vec{r}_s$ and carries charge q_s, the electric field it creates at any position P is

$$\vec{E}_s(\mathrm{P}) = k\frac{q_s}{r_{sP}^2}\hat{r}_{sP}, \tag{23.4}$$

where $\hat{r}_{sP}$ is a unit vector pointing from the source particle to P and $k = 9.0 \times 10^9\ \mathrm{N \cdot m^2/C^2}$.

Principle of **superposition of electric fields:** The combined electric field $\vec{E}$ created by a system of particles that carry charges $q_1, q_2, \ldots$ is

$$\vec{E} = \vec{E}_1 + \vec{E}_2 + \cdots = \Sigma k\frac{q_i\hat{r}_{iP}}{r_{iP}^2}. \tag{23.5}$$

Electric dipoles (Sections 23.4, 23.6, 23.8)

Concepts A *uniform electric field* has a constant magnitude and direction at all locations; a *nonuniform electric field* can vary in magnitude and/or direction from one location to another. A uniform electric field gives a charged particle a constant acceleration.

An **electric dipole** consists of equal quantities q_p of positive and negative charge separated by a small distance.

When an electrically neutral object is placed in an electric field, the field causes the object's positive and negative charge centers to separate, producing an **induced dipole.**

Quantitative Tools The **dipole moment** $\vec{p}$ (C · m) of a dipole is

$$\vec{p} \equiv q_p\vec{r}_p, \tag{23.9}$$

where $\vec{r}_p$ is the vector from the center of negative charge to the center of positive charge.

The torque $\Sigma\vec{\tau}$ produced on a dipole by a uniform electric field is

$$\Sigma\vec{\tau} = \vec{p} \times \vec{E}. \tag{23.21}$$

The **induced dipole moment** $\vec{p}_{ind}$ of a neutral atom placed in an external electric field that is not too large is

$$\vec{p}_{ind} = \alpha\vec{E}, \tag{23.24}$$

where α is the **polarizability** of the atom.

For positions far from an electric dipole, the electric field due to the dipole is inversely proportional to the cube of the distance r from the dipole. In particular, at distant points along the bisector axis (x axis)

$$E_y \approx -k\frac{p}{|x^3|}, \tag{23.10}$$

and at distant points along the dipole separation axis ($+y$ axis)

$$E_y \approx 2k\frac{p}{y^3}. \tag{23.13}$$

Electric fields of continuous charge distributions (Section 23.7)

Concepts **Charge density** can be linear (λ, C/m), over a surface area (σ, C/m^2), or throughout a volume (ρ, C/m^3).

Quantitative Tools The electric field due to a set of infinitesimal source particles of charge dq_s is

$$\vec{E} = k\int \frac{dq_s}{r_{sP}^2}\hat{r}_{sP}. \tag{23.15}$$

See the Procedure box on page 828.

Linear, *area*, and *volume charge densities* are

$$\lambda \equiv \frac{q}{\ell} \tag{23.16}$$

$$\sigma \equiv \frac{q}{A} \tag{23.17}$$

$$\rho \equiv \frac{q}{V}. \tag{23.18}$$

A uniformly charged rod of length ℓ carrying charge q oriented along the y axis and centered at the origin produces an electric field

$$E_x = \frac{kq}{x\sqrt{\ell^2/4 + x^2}}$$

along the x axis.

A uniformly charged ring of charge q and radius R oriented with the plane of the ring perpendicular to the z axis and the ring center at the origin produces an electric field

$$E_z = k\frac{qz}{(z^2 + R^2)^{3/2}}$$

along the z axis.

A uniformly charged disk of surface charge density σ and radius R oriented with the disk face perpendicular to the z axis and the disk center at the origin produces an electric field

$$E_z = 2k\pi\sigma\left[1 - \frac{z}{(z^2 + R^2)^{1/2}}\right]$$

along the z axis.

Outside a solid sphere of radius R carrying a uniformly distributed charge q, the electric field produced by the sphere is

$$E_{\text{sphere}} = k\frac{q}{r^2},$$

where $r > R$ is the distance from the sphere's center. This field is the same as the field that would be produced if all the charge were concentrated at the sphere's center.

Questions and Problems

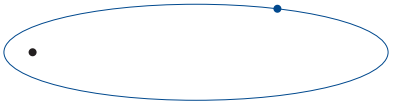
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

23.1 The field model

1. Compare the magnitudes of the gravitational field that earth feels due to the Sun and the Moon. •
2. Determine the magnitude of Earth's gravitational field for (*a*) a 70.0-kg person standing at Earth's surface, (*b*) a 700.0-kg satellite in orbit 150 km above Earth's surface, and (*c*) the Moon (use 3.844×10^8 m for the center-to-center distance between Earth and the Moon). •
3. Sketch the gravitational field at 15 positions uniformly distributed along the comet orbit in Figure P23.3. ••

Figure P23.3



4. Suppose there is an electric field pointing horizontally toward the east at some location near Earth's surface. Does it make sense to add the gravitational and electric *fields* to determine the behavior of a proton at that location? ••

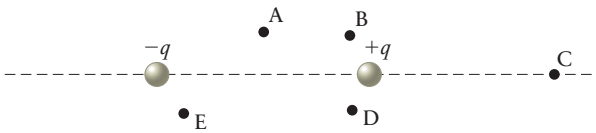
23.2 Electric field diagrams

5. The units of the gravitational field are those of acceleration. Is that true of the electric field as well? If not, why not? •
6. Assume an electron is fixed at the origin of a rectangular Cartesian coordinate system. Calculate the magnitude of the electric field at 15 mm on the positive *x*-axis. •
7. The dwarf planet Pluto and its moon Charon have very similar masses. Suppose the masses are equal, and draw a gravitational field diagram for this system of two source objects. ••
8. Draw a vector field diagram for particles carrying charges $+2q$ and $-q$ separated by a distance r. Comment on the significance of the vector diagram. ••

23.3 Superposition of electric fields

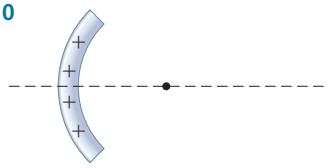
9. The two Ping-Pong balls in Figure P23.9 carry charges that have the same magnitude but opposite signs. Sketch the vectors that represent the electric fields the balls produce at locations A, B, C, D, and E, which all lie in a common plane that includes the line joining the centers of the balls. Use a ruler to estimate that relative distances between locations. •

Figure P23.9



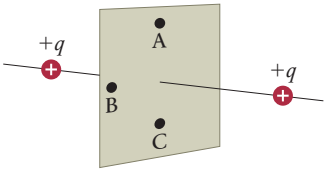
10. Positive charge is distributed uniformly on a plastic rod bent to form a quarter-circular arc (Figure P23.10). What is the direction of the electric field at the center of the circle? •

Figure P23.10



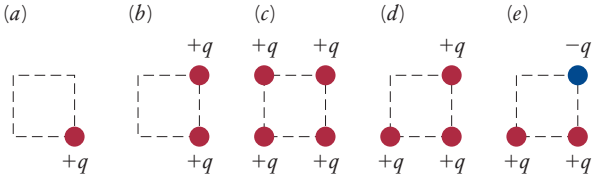
11. Repeat Problem 9 for the case where the positive charge in Figure P23.9 remains $+q$ but the negative charge is changed to $-2q$. ••
12. The two particles in Figure P23.12 carry identical charges. Sketch the vectors that represent the electric fields the particles produce at locations A, B, and C, which all lie in the plane that is midway between the two particles and perpendicular to the line connecting them. ••

Figure P23.12



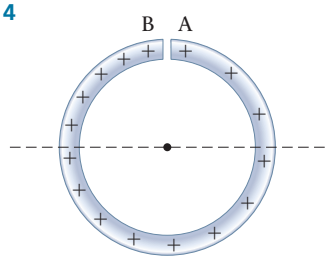
13. Charged beads are placed at the corners of a square in the various configurations shown in Figure P23.13. Each red bead carries a charge $+q$, and the blue bead carries a charge $-q$. Rank the configurations according to the magnitude of the electric field at the center of the square, smallest magnitude first. ••

Figure P23.13



14. The charge on a nonconducting rod increases linearly from end A to end B. The rod is bent in a circle so that ends A and B almost meet very near the top of the circle (Figure P23.14). What is the direction of the electric field at the center of the circle? (Hint: Consider contributions from diametrically opposite segments!) •••

Figure P23.14



23.4 Electric fields and forces

15. The Burj Khalifa in Dubai has 163 floors. Compare the gravitational field in a room on the ground floor and one on the 163rd floor. •
16. A positively charged test particle is placed midway between two fixed, identical positively charged source particles. (*a*) Is the test particle in a stable or unstable equilibrium at that location? (*b*) If the test particle is replaced by a negatively charged test particle, is it in a stable or unstable equilibrium at that location? •
17. A box full of charged plastic balls sits on a table. The electric force exerted on a ball near one upper corner of the box has components 1.2×10^{-3} N directed north, 5.7×10^{-4} N directed east, and 2.2×10^{-4} N directed vertically upward. The charge on this ball is 120 nC. If this ball were replaced with one that has a charge of −50 nC, what would be the force components exerted on the replacement ball? ••
18. An electron traveling horizontally east passes between two horizontal, oppositely charged plates and is deflected downward. Passing through the same space between the plates, in which direction (if any) would each of the following be deflected: (*a*) a proton traveling horizontally east, (*b*) an electron traveling horizontally west, (*c*) a proton traveling horizontally west, and (*d*) a proton traveling horizontally north? ••
19. A 10.0-mg oil drop carrying a charge of +15 μC passes undeflected through a region in which there is a uniform, constant electric field. What is the magnitude and direction of the electric field? ••
20. Each dipole in Figure P23.20 is free to rotate about an axis that is perpendicular to the plane of the page (represented by the dot in the center of each dipole). The dipoles are initially held fixed in a horizontal line as shown but are then released and begin rotating. What is their most likely final orientation if there is a mechanism for energy dissipation? ••

Figure P23.20



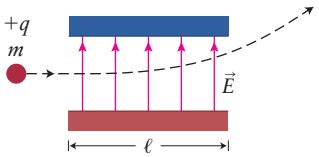
21. Suppose two plates lie in parallel horizontal planes, one plate in the *xy* plane at $z = 0$ and the other plate in the plane that is parallel to the *xy* plane at $z = 10$ mm. Between the plates is a constant electric field directed vertically upward (that is, in the positive *z* direction). A proton and an electron are launched in the positive *x* direction with the same initial velocity from position (0, 0, 5.0 mm). If the proton strikes a plate at (200 mm, 0, 10 mm), where does the electron strike? •••

23.5 Electric field of a charged particle

22. Two uniformly charged pellets A and B are held some distance from each other, and then the charge on A is doubled. Which of the following statements is most correct? (*a*) The magnitude of the electric force exerted on A is doubled because the electric field at the position of A is doubled. (*b*) The magnitude of the electric force exerted on A is doubled because the electric field at the position of B is doubled. (*c*) The magnitude of the electric force exerted on B is doubled because the electric field at the position of B is doubled. •
23. What is the magnitude the net force on an electron when placed at the electric field at a distance of 100 mm produced by 5 μC of charge? •
24. A small, charged, spherical object at the origin of a Cartesian coordinate system contains 3.30×10^4 more electrons than protons. Find the magnitude and direction of the electric field at a point that is 3 mm away from the origin. •
25. Consider a hydrogen atom. The Bohr radius of hydrogen atom is 0.529×10^{-8} cm. Calculate the electric field halfway between the electron and the proton. •
26. A positively charged particle initially at rest on the ground moves 4.0 m upward in 2.00 s. If the particle has a charge-to-mass ratio of 10 μC/g and the electric field in this region is constant and uniform, what is the magnitude and direction of the electric field? ••
27. A particle, carrying a positive charge of 4 nC, located at (5cm, 0) on the *x*-axis experiences an attractive force of magnitude 115.2 N due to an unknown charge *q* located at the origin. (*a*) Determine the value of *q*. (*b*) Using the relationship between electric field and electric force and the result from part (*a*), determine the magnitude and direction of the electric field due to the charge *q* at a point located 20 cm away from the 4-nC charge in a direction perpendicular to the line between the two charges. ••
28. A particle at the origin of a Cartesian coordinate system carries a charge of 5.0×10^{-9} C. What are the magnitude and direction of the electric field at (*a*) (4.00 mm, 0), (*b*) (0, 4.00 mm), and (*c*) (−3.00 mm, 4.00 mm)? ••
29. (*a*) Plot the value of the radial component of the electric field due to a positively charged pellet as a function of distance from the pellet center. (*b*) Repeat for a negatively charged pellet. Take each pellet to be a uniformly charged nonconducting sphere having a radius of 18.5 mm and carrying a charge of 1.11 nC. ••
30. A rubber flying ring having a diameter of 20 cm is fixed to the *xy*-plane of the Cartesian coordinate system such that its geometric center lies on the origin. The ring is uniformly charged with +10 μC. (*a*) What is the magnitude of the electric field at 10 cm on the z-axis? (*b*) What is the magnitude of the electric field at the origin? ••
31. A small sphere carrying 6.0 nC of charge is placed 100 mm from a small sphere carrying 3.0 nC of charge. Assume the spheres are tiny relative to the 100-mm separation distance. At what position on the line joining the spheres is the electric field magnitude zero? ••
32. Two beads, one carrying charge +*q* and the other carrying charge +4*q*, are separated by a distance *d* that is much greater than the radius of each bead. (*a*) Is there any location along the line between them where the electric field magnitude is zero? If so, what is that location in terms of *d*? (*b*) Is there any other position where the electric field magnitude is zero? ••
33. In an inkjet printer, tiny drops of ink of inertia *m* are given a charge *q* and then fired toward the paper at speed *v*. They first pass between two charged plates of length ℓ that create a uniform electric field of magnitude *E* between

them (Figure P23.33). The field direction is perpendicular to the plates and to the initial path of the drops. The electric field deflects each drop from its initial horizontal path and is used to control what is printed on the paper. (*a*) Derive an expression for how far in the vertical direction a drop is deflected as it passes between the plates. (*b*) For a drop inertia of 1.5×10^{-10} kg, an electric field strength of 1.2×10^{6} N/C, a plate length of 10 mm, and a drop speed of 20 m/s, calculate the drop charge necessary to result in a deflection of 1.3 mm. (*c*) Is it acceptable to ignore the effects of gravity in this case? ••

Figure P23.33



34. (*a*) What electric field magnitude is needed to balance the gravitational force exerted by Earth on an electron near Earth's surface? (*b*) Relative to the electron's position, where would you have to place a proton in order to create an equivalent electric force on the electron? ••

35. Consider a rectangle with diagonal length $2a$ in the yz plane with the origin at the center of the rectangle. Four beads, each carrying charge q, are placed on the rectangle perimeter, one bead at each corner. Show that the electric field along the x axis is given by

$$E_x = k\left(\frac{4qx}{(x^2 + a^2)^{3/2}}\right)\hat{\imath}. \text{ ••}$$

36. A particle carrying a charge of 6.0 μC is located at the origin of a rectangular coordinate system, and a particle carrying a charge of 4.0 μC is located at (0, 5.0 m). What are the magnitude and direction of the electric field at the positions (*a*) (5.0 m, 0), (*b*) (−5.0 m, 0), and (*c*) (0, −5.0 m)? ••

37. Five charged particles are located around a semicircular arc of radius 100 mm, with one particle at each end of the semicircle and the remaining three spaced equally between the two ends. The semicircle lies in the region $x < 0$ of an xy plane, such that the complete circle is centered on the origin. (*a*) If each particle carries a charge of 1.00 nC, what is the electric field at the origin? (*b*) Where could you put a single particle carrying a charge of −5.00 nC to make the electric field magnitude zero at the origin? ••

38. A positively charged particle 1 is at the origin of a Cartesian coordinate system, and there are no other charged objects nearby. You need the electric field magnitude at the position (3.00 nm, 4.00 nm) to be zero, so that no electric force is exerted on any charged object placed at that location. (*a*) Make a plot of all positions where a single additional charged particle 2 (of arbitrary charge) could be placed to make the electric field magnitude zero at (3.00 nm, 4.00 nm). Explain some of the key points of your graph. (*b*) Where would you need to place particle 2 if it carried the same charge as particle 1? ••

39. Three particles carrying equal positive charge q are located at the corners of an equilateral triangle with side length a. What are the magnitude and direction of the electric field (*a*) at the center of the triangle, (*b*) at the midpoint of any side of the triangle, and (*c*) a distance a above the triangle apex? •••

40. Two nonconducting spheres 1 and 2 carry the same charge, and the magnitude of the electric force exerted by each sphere on the other is 0.10 N when they are 50 mm apart. (*a*) What is the magnitude of the charge on each sphere, assuming each has a diameter much smaller than 10 mm? (*b*) What is the magnitude of the electric field 50 mm directly above sphere 1 if sphere 2 is located to the left of sphere 1? (*c*) Would your answers change if each sphere had a radius of 10 mm? •••

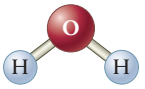
41. Three particles carrying charges $2q$, q, and $-3q$ are located at the vertices of an equilateral triangle on the xy-plane. Particles 1 and 2 are located symmetrically on the x-axis and particle 3 is located on the positive y-axis. Find the direction of the electric field at the center of the triangle. (Hint: You need to use only ratios and geometrical considerations.) •••

42. In table salt—sodium chloride, NaCl—the Na^+ and Cl^- ions are arranged in a cubic crystal structure. You can observe in this structure a cube made up of eight ions: four Na^+ and four Cl^- in alternating corners. Each Na^+ has a radius of 99 pm, each Cl^- has a radius of 181 pm, and the ions are spaced as though they were solid balls of these sizes, touching their neighbors. (*a*) For a single cube of this arrangement, what are the magnitude and direction of the electric field at one of the Na^+ ions due to the other seven ions in the cube? (*b*) What are the magnitude and direction of the vector sum of the electric forces exerted on this Na^+ ion by the other seven ions? •••

23.6 Dipole field

43. A water molecule (Figure P23.43) has a dipole moment of 6.19×10^{-30} C·m. If the O–H bonds were ionic bonds (in reality they are polar covalent bonds), the electrons from the two hydrogen atoms would be completely transferred to the oxygen atom. Using the measured dipole moment and the fact that the centers of positive and negative charge are 0.058 nm apart, can you rule out the hypothesis that the O–H bonds are ionic? •

Figure P23.43



44. Two balls of identical size (ball 1 made up of nylon and ball 2 made up of rubber) were rubbed against each other so that they carry a uniformly distributed charge of magnitude 0.25 nC. After charging, the two balls were stuck to the end of a 15 cm scale. In that situation, calculate the dipole moment of the arrangement. •

45. As Figure P23.43 shows, the water molecule is bent, and the angle formed by the three atoms is 104.5°. Given that the permanent dipole moment of the molecule is 6.186×10^{-30} C·m, determine the dipole moment of each O–H bond. ••

46. A proton located several proton diameters away from a small charged object carrying charge q is subject to an

electric field of magnitude E. As the proton moves a distance d along the x axis away from the object, the electric field magnitude drops to $E/4$. If the charged object had instead been an electric dipole (with a charge of magnitude q on each end) oriented in the z direction, by what factor would the initial electric field magnitude E have dropped as the proton receded by distance d along the x axis? ••

47. Two ping pong balls, 1 and 2, are charged uniformly such that ball 1 gets 0.5 nC charge and ball 2 gets −0.5 nC charge. Calculate the ratio of the magnitudes of the dipole moments produced if they were stuck to the opposite ends of the two sticks of length 200 mm and 400 mm. ••

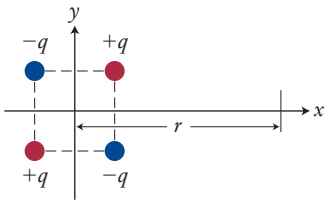
48. Equation 23.12,

$$E_y = k\frac{q_p}{y^2}\left[\left(1-\frac{d}{2y}\right)^{-2}-\left(1+\frac{d}{2y}\right)^{-2}\right],$$

was derived for the case $y > d/2$. Explain why this equation holds for the case $y < -d/2$. ••

49. An electric *quadrupole* can be constructed by placing four charged objects at the corners of a square (Figure P23.49). The objects are identical except for the charge they carry: The two objects in one diagonally opposed pair each carry charge $+q$, and the other two objects carry charge $-q$. (Note that there is no charge on the quadrupole overall and that the combination of four charges has no dipole moment.) How does the magnitude of the electric field depend on the distance r from the center of the quadrupole? Take the side length for the square to be d and assume $r \gg d$. ••

Figure P23.49



50. A dipole is centered at the origin of a coordinate system, and a small charged sphere is some distance away along the perpendicular bisector of the dipole. The particle carries a uniformly distributed charge of −3.0 nC, and experiences a 200 nN electric force in the positive y direction. (*a*) If the dipole charge is 10 nC and the dipole separation is 20 mm, how far away from the dipole is the sphere? (*b*) How is the dipole moment oriented? ••

51. Two thin plastic rods, each of length ℓ, are joined end to end. One rod is positively charged with a uniform linear charge density λ, and the other carries a negative charge density $-\lambda$. What is the effective dipole moment of the rod? ••

52. An electric dipole that has dipole separation d is aligned along the y axis of an xy coordinate system, pointed in the positive y direction. (*a*) Show that for x and y much greater than d, the x and y components of the electric field of the dipole are given by

$$E_x = \frac{3kpxy}{(x^2+y^2)^{5/2}};\quad E_y = \frac{kp(2y^2-x^2)}{(x^2+y^2)^{5/2}}.$$

(*b*) Show that this general result includes the special results derived in Section 23.6. •••

23.7 Electric fields of continuous charge distributions

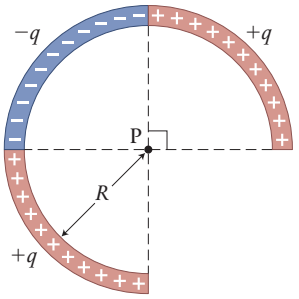
53. You've been given the task of charging a spherical weather balloon made of a conducting material, and you want to put as much charge on it as possible. An experienced colleague advises you that the air surrounding the balloon becomes conducting when the electric field at the balloon surface reaches 100,000 N/C, causing a spark that discharges the balloon. Because the balloon must stay charged, you cannot allow this air breakdown to occur. By measuring the balloon's shadow, you calculate the balloon diameter to be 3.5 m. This gives you enough information to estimate the maximum charge the balloon can have without causing air breakdown. •

54. A positively charged particle is released from rest along the axis of symmetry of a fixed ring carrying a uniformly distributed negative charge. Describe the motion of the particle. •

55. A 20-cm-long rod, with uniform linear charge density 100 nC/cm, is set up symmetrically on the x axis. What are the magnitude and direction of the electric field at a point located at a distance of 5cm on the perpendicular bisector of the rod in the xy-plane? •

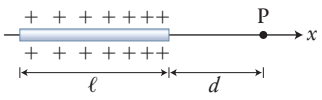
56. In the uniform charge distribution shown in Figure P23.56, each of the three arcs forms one-fourth of the circumference of a ring. The upper right and lower left arcs each carry a positive charge q, while the upper left arc carries a charge $-q$. Determine the electric field at P, the ring center, in terms of q and R. ••

Figure P23.56



57. The thin glass rod of length ℓ in Figure P23.57 has a linear charge density that starts out as zero at the left end of the rod and increases linearly from left to right. The positive charge on the rod is q_{rod}. (*a*) What is the electric field along the rod's axis at position P, which lies a distance d from the right end of the rod? (*b*) What is the approximate expression for the magnitude of the electric field at distances $d \gg \ell$—that is, far enough away to make the rod look small? ••

Figure P23.57



58. You have two disks, 1 and 2, both of radius $R = 25.0$ mm and both made of the same nonconducting material. Disk 1 carries a uniformly distributed charge $q_1 = 1.50\ \mu$C, and the uniformly distributed charge q_2 on disk 2 is unknown. You place disk 1 at $z_1 = 0$ and disk 2 at $z_2 = +32.0$ mm such that the symmetry axis of each disk coincides with the z axis. Using a small charged pith ball, you locate one position on the z axis where the electric force exerted on the pith ball is zero; this position is $z_{E=0} = +88.0$ mm. What is the charge on disk 2? ••

59. What percentage error in the electric field magnitude do you introduce by approximating the charged disk of Figure 23.29 as an infinite charged sheet with the same surface charge density for (*a*) $z = 0.1R$, (*b*) $z = 0.5R$, and (*c*) $z = R$? ••

60. You wish to determine the electric field magnitude along the perpendicular bisector of a 250-mm line along which 30 nC of charge is distributed uniformly. You want to get by with a minimal amount of work, so you need to know when it is sufficient to approximate the line of charge as a charged particle. At what distance along the perpendicular bisector does your error in E reach 5% when you use this approximation? ••

61. For a uniformly charged disk of radius 90 mm, you wish to determine the electric field magnitude along the axis that runs through the disk center and perpendicular to the disk face. You want to get by with a minimal amount of work, so you need to know when you can get by with approximating the disk as an infinite sheet. At what distance along the perpendicular axis does your error in E reach 10% when you use this approximation? ••

62. Three narrow concentric rings of radii 50 mm, 70 mm, and 90 mm are centered about the origin, with the axis of symmetry of each ring oriented along the y axis. The charge on the inner ring is 1.0 μC, that on the middle ring is $-2.0\ \mu$C, and that on the outer ring is 1.0 μC. (*a*) What is the electric field magnitude at (0, 100 mm, 0)? (*b*) What quantity of charge must be placed at (0, −100 mm, 0) to make the electric field magnitude zero at the position calculated in part *a*? ••

63. Two large oppositely and uniformly charged parallel plates are held 10 mm apart. An electron is projected halfway between the plates and parallel to them with an initial speed of 5×10^6 m/s. The electron hits the top plate 35 mm from where it was projected. (*a*) Which plate is negatively charged? (*b*) What are the surface charge densities on the plates? ••

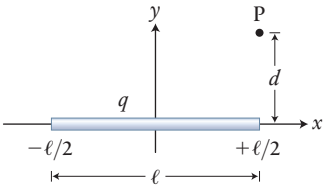
64. A particle of inertia m that carries charge q is held above an infinite sheet on which the surface charge density is σ. The charge on the sheet is of the same type as the charge on the particle. The particle is then released from rest. (*a*) Describe its motion. (*b*) What is its acceleration? (*c*) What is the change in its kinetic energy after it has traveled a distance s? ••

65. A uniformly charged thin rod lies along the x axis from $x = 0$ to $x = +\infty$. (*a*) Derive an expression for the component E_y of the electric field in the positive y direction, and (*b*) show that the electric field at all positions on this axis makes an angle of 135° with the rod. •••

66. One half of a charged thin ring of radius R carries a charge q_1 uniformly distributed over it, and the other half of the ring carries a charge q_2 uniformly distributed over it. For any position along the axis of symmetry of the ring, derive an expression for (*a*) the component of the electric field parallel to the axis and (*b*) the component of the electric field perpendicular to the axis. •••

67. In Figure P23.67, determine the x and y components of the electric field at position P, which is a distance d above one end of a rod of length ℓ carrying uniformly distributed charge q. •••

Figure P23.67



68. A thin rod of length 2ℓ has a linear charge density that is λ_0 at the left end but decreases linearly with distance going from left to right in such a way that the charge on the entire rod is zero. (*a*) What is the magnitude of the electric field along the rod's axis at a position P that is a distance $d > \ell$ to the right of the center of the rod? (*b*) Does your expression simplify to a familiar result for the limiting case $d \gg \ell$? (Hint: The dipole moment of the rod has magnitude $p = 2\lambda_0\ell^2/3$.) •••

69. You place 2.2 μC of charge along a long nonconducting rod. The rod extends from $x = 0$ farther than you can see along the positive x axis. The charge distribution has the form of a decreasing exponential: $q(x) = q_0 e^{-x/\ell}$, where $\ell = 28.6$ mm. Calculate the electric field magnitude 20 mm from the near end of the rod, along its long axis. •••

23.8 Dipoles in electric fields

70. The water molecule is a permanent dipole with a dipole moment of 6.186×10^{-30} C·m. If a single water molecule were oriented such that its dipole moment is along the z axis, what torque is caused on it by the electric force due to an 8500-N/C electric field if that field is in the xz plane and directed 42° above the x axis? •

71. A microwave oven, which fills the oven chamber with oscillating electric fields, works well at heating food that has a high water content. However, it does not work well with frozen food or with food that has a high oil content but low water content. Can you think of a reason for this? (Hint: Water molecules in ice are much more rigidly constrained than they are in liquid, and oil molecules are nonpolar.) •

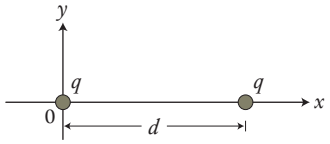
72. The electric force due to a uniform external electric field causes a torque of 8×10^{-9} N·m on an electric dipole oriented at 45° from the direction of the external field. The dipole moment of the dipole is 4.0×10^{-12} C·m. (*a*) What is the magnitude of the external electric field? (*b*) If the two particles that make up the dipole are 5 mm apart, what is the charge on each particle? •

73. Among copper and gold, whose polarizability is higher? Justify your answer. •
74. Two oppositely charged identical metallic spheres, each having a mass of 4 g, lie on a horizontal plane, joined by a tiny insulating rod of length 2 cm. The magnitude of the charge on each sphere is 3 μC. The rod is parallel to a uniform electric field of magnitude 1500 N/C. Calculate the angular frequency ω of small oscillations about the equilibrium orientation. ••
75. A dipole that is free to move is placed near a fixed dipole, with the midpoint of the free dipole on the perpendicular bisector of the fixed dipole. The distance from the free dipole to the fixed one is much greater than the dipole separation d of the fixed dipole. Draw a diagram to illustrate the motion of the free dipole if it is placed with its dipole moment (*a*) in the same direction as the direction of $\vec{p}_{\text{fixed}}$, the dipole moment of the fixed dipole, (*b*) in the direction opposite that of $\vec{p}_{\text{fixed}}$, and (*c*) perpendicular to the direction of $\vec{p}_{\text{fixed}}$. ••
76. In a particular region of space, an electric field has a constant direction but its magnitude increases smoothly along that direction. A dipole is placed in this field, its dipole moment is oriented perpendicular to the field, and it is released from rest. If no dissipative forces are present, describe the dipole's motion, paying particular attention to how it differs from what the motion would be if the electric field were uniform. ••
77. A particle carrying 5.0×10^{-7} C of charge is located on the perpendicular bisector of a small dipole, 300 mm from the center of the line joining the two poles of the dipole. The magnitude of the electric force exerted on the particle is 10.0×10^{-6} N. Draw a diagram showing the dipole and indicating the direction of its dipole moment. On this diagram, show (*a*) the direction of the electric force exerted on the particle and (*b*) the direction of the electric force exerted on the dipole. (*c*) Determine the magnitude of the electric force exerted on the dipole. ••
78. There are two possible alignments of a dipole in an external electric field where the dipole is in equilibrium. (*a*) Draw a field diagram for a uniform electric field and show these two alignments. (*b*) Are both alignments stable? (Consider what would happen in each case if you gave the dipole a slight twist.) (*c*) Based on your answer to part *b* and your experience in mechanics, in which orientation does the dipole have less potential energy? ••
79. In very rare cases, the polarizability of a molecule may be negative. (*a*) Draw a picture showing the resultant charge distribution when such a molecule is placed in the electric field surrounding a positively charged particle. (*b*) Describe the motion of the molecule that would result. ••
80. Two identical particles, each carrying charge q, are a distance r apart. A dipole for which the dipole moment magnitude is p and the dipole separation is $d < r$ is placed on the (imaginary) line joining the two particles in such a way that the center of the dipole coincides with the center of the line joining the two particles. Determine the electric force exerted on the dipole. (Hint: You may wish to use a Taylor series expansion to determine the small difference in the electric forces exerted on the two poles of the dipole.) •••
81. An electrically neutral molecule is collinear with (and located between) two charged particles, one carrying a charge of +3.56 μC and the other carrying a charge of −1.05 μC. The center of the molecule is 2.57 μm from each particle. If the vector sum of the electric forces exerted on the molecule is 45.0 nN, what is the polarizability of the molecule? •••

Additional Problems

82. A plastic flying disc with radius 25 cm is uniformly charged with $\sigma = 8$ nC/m^2. Find the magnitude of the electric field at 30 m on the axial position. •
83. Four particles are located at the corners of a square that is 50 mm on a side. All four particles carry a charge of magnitude 3.0 nC, with positive charge at the lower left corner and negative charge at the other corners. Draw the electric field vectors at the center of the square caused by each of the four source charges, then determine the magnitude and direction of the electric field at the center of the square. •
84. A charge of 2.86 μC is uniformly distributed within a rubber ball having radius 5.6 cm. Calculate the electric field at 10 cm from the center of the rubber ball. •
85. Each of the two particles in Figure P23.85 carries a charge of magnitude q. Determine a position, if one exists, where the electric field magnitude is zero if the charge (*a*) is of the same type on the two particles and (*b*) is positive on one particle and negative on the other. •

Figure P23.85



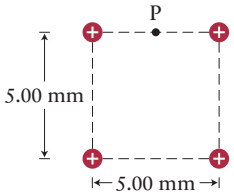
86. Two uniformly charged plates carrying opposite charges are separated by a distance of 20 mm. The magnitude of the charge density on each plate is 3×10^{-8} C/m^2. An electron is launched with a speed of 2.1×10^6 m/s halfway between the two plates. (*a*) How far does the electron travel before it hits one of the plates? (*b*) What speed does the electron attain before it hits a plate? ••
87. An electrically neutral, linear, polar molecule has positive charge equivalent to the charge on one proton centered 0.30 nm from its center of mass and an equal quantity of negative charge centered 0.10 nm from its center of mass (opposite the positive charge). The two charge centers are collinear with the molecule's center of mass. In an electric field of magnitude 100.0 N/C, what are (*a*) the greatest possible and (*b*) the smallest possible torque experienced by this molecule? ••
88. You have a pair of objects that interact with each other electrically. The objects initially are separated by a distance r and exert a force of magnitude F^{E} on each other. If the separation is increased to $2r$, calculate the new force magnitude between the objects if they are (*a*) charged particles, (*b*) permanent dipoles with their separation along their dipole axis, (*c*) permanent dipoles

with their separation perpendicular to their dipole axis, (*d*) a charged particle and a permanent dipole with the dipole pointing directly away from the charged particle, (*e*) a charged particle and an induced dipole. ••

89. Two infinitely wide charged plates carry opposite charge densities of magnitude $1.77 \times 10^{-6}\ \text{C/m}^2$. A small sphere, of mass m = 2 g, lies between the plates, attached to the positive charged plate by means of an inextensible insulating wire of negligible mass. At equilibrium, the wire forms an angle of 30° with respect to the positively charged plate. Find the magnitude of the charge on the sphere. ••

90. Figure P23.90 shows four charged particles, each having a charge equal to +3.00 μC, arranged in a square. What are the magnitude and direction of the electric field at P, located midway between the top two particles? ••

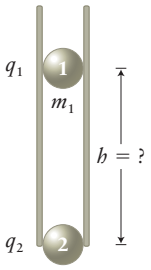
Figure P23.90



91. A uniformly charged rod extends from $y = -150$ mm to $y = +150$ mm along the y axis of an xy coordinate system. The charge on the rod is 30 nC. (*a*) With the goal of eventually calculating the electric field magnitude along the rod's perpendicular bisector, approximate the rod as three charged particles, and draw a picture showing the locations of the three particles you will use for the calculation and the magnitude of the charge on each one. (*b*) What is the electric field magnitude at $x = 200$ mm using this approximation? (*c*) What is the error in your result in part *b*? (*d*) If you were unable to integrate, what could you do to your approximation to make it better? ••

92. A small sphere 1 carrying charge q_1 and having inertia m_1 is constrained to moving inside a narrow vertical tube (Figure P23.92). Fixed at the bottom of the tube is a small sphere 2 carrying charge q_2. (*a*) Determine the equilibrium height h for sphere 1 (ignore friction). (*b*) Show that if sphere 1 is displaced from this equilibrium position by a small amount and then released, it moves in simple harmonic motion with angular frequency $\omega = (2g/h)^{1/2}$. •••

Figure P23.92



93. Your boss is designing a new data storage system that will have a lot of dipoles very close to one another. They will point in the z direction and be closely spaced in the x and y directions. The data reader will detect E_z as it passes some small distance above the sheet of dipoles. You know that Equation 23.13 is valid only for large distances from the dipole, but it is much easier to apply than the exact result. You prefer the easiest calculation possible, but your boss demands 1.00% accuracy at distances only a few dipole separations from one of the source dipoles. ••• CR

94. You are designing a new guidance mechanism for an inkjet printer. You believe that a circular arc of charge, mounted in the xy plane, is needed in order to deflect the ink droplets most efficiently from their original path along the z axis. You want to use a uniformly charged arc of arc length s that subtends an angle 2θ along a circle of radius R through whose center the ink droplets will pass. The amount of charge that can easily be transferred to the arc is q, which limits the magnitude of the electric field that can be attained at the center of the circle. ••• CR

95. A rod of length πR is composed of three nonconducting segments of equal length. The middle segment is electrically neutral, and each end segment carries a uniformly distributed negative charge $-q$. The rod is bent into a semicircle of radius R. Determine the magnitude of the electric field at the center of the semicircular arc. •••



24

Gauss's Law

24.1 Electric field lines
24.2 Field line density
24.3 Closed surfaces
24.4 Symmetry and Gaussian surfaces
24.5 Charged conducting objects

24.6 Electric flux
24.7 Deriving Gauss's law
24.8 Applying Gauss's law

CONCEPTS

QUANTITATIVE TOOLS

In principle, Coulomb's law allows us to calculate the electric field produced by any discrete or continuous distribution of charged objects. In practice, however, the calculation is often so complicated that the sums or integrals that arise might require numerical evaluation on a computer. For this reason, it pays to search for additional methods to determine the electric field produced by a charge distribution. In this chapter we develop a relationship between an electric field and its source, known as *Gauss's law*, that can be used to determine the electric fields due to charge distributions that exhibit certain simple symmetries. These symmetries appear in many common applications, which makes Gauss's law an important tool in calculating electric fields. As we shall see in Chapter 30, Gauss's law is one of the fundamental equations of *electromagnetism*—the theory that describes electromagnetic interactions and electromagnetic waves.

24.1 Electric field lines

In Chapter 23 we used vector field diagrams to visualize electric fields. Another way to visualize electric fields, which will help us reach some new insights, is to draw **electric field lines.** These lines are drawn so that at any location the electric field $\vec{E}$ is tangent to them. Because the electric field is a vector, we assign to field lines a direction that corresponds with the direction of the electric field.

To draw an electric field line, imagine placing a test particle carrying a positive charge q_t somewhere near a charge distribution. Then move the test particle only a small distance in the direction of the electric force exerted on it. (Remember from Chapter 23 that the electric field points in the same direction as the electric force exerted on a positively charged test particle.) Slowly repeat the procedure to trace out a line (**Figure 24.1**). We label the field lines with the symbol E to remind us that they represent an electric field.

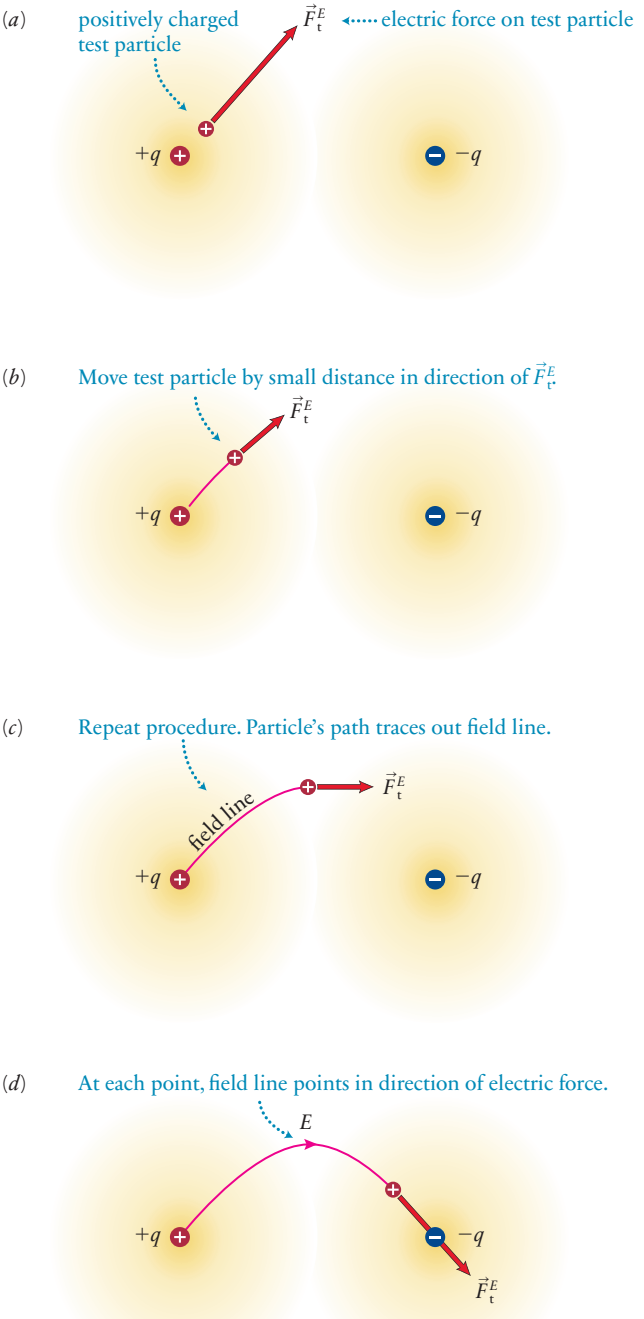
CONCEPTS

24.1 Draw several field lines representing the electric field of an isolated positively charged particle. Repeat for a negatively charged particle.

As Checkpoint 24.1 illustrates, the field line diagrams for a positive and for a negative isolated charged particle are similar, even though the electric fields point in opposite directions. They point radially outward from a positively charged particle and point radially inward toward a negatively charged particle. This direction means that electric field lines always start from a positively charged object and always end on a negatively charged object, never the other way around.

Because an electric field is present everywhere around a charged object, a field line passes through every location in space. In practice, we draw only a finite number of field lines to represent the entire field. **Figure 24.2**, for example, shows the pattern of field lines created by a pair of oppositely charged particles (that is, a dipole).

Figure 24.1 Using a positively charged test particle to trace out an electric field line.



Sixteen field lines emanate from the positively charged particle on the left, and 16 field lines terminate on the negatively charged particle on the right. Notice the correspondence between this pattern and the corresponding vector field diagram in **Figure 24.3*a*** and the pattern created by the fibers in Figure 24.3*b*.

The number of field lines that emanate from a positively charged object is arbitrary; we could have chosen some number other than 16 for Figure 24.2. However, in a given field line diagram, the number of

Figure 24.2 Electric field line diagram for an electric dipole.

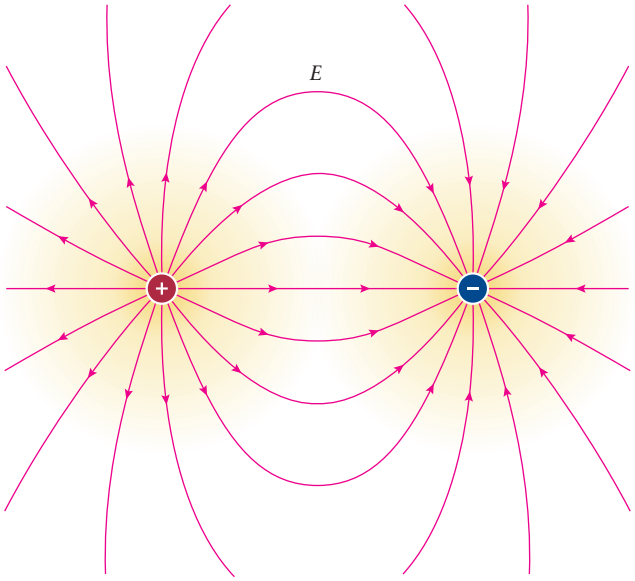
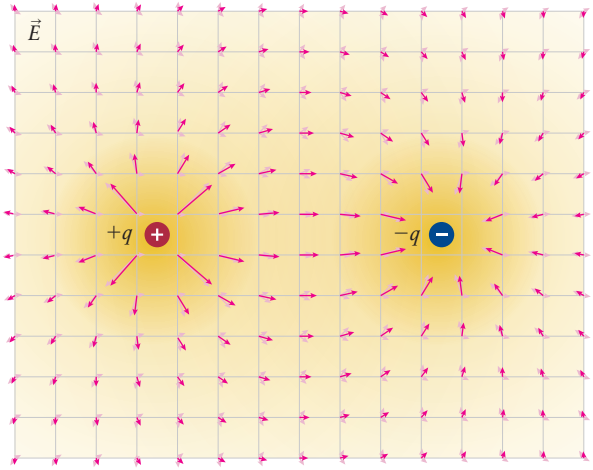
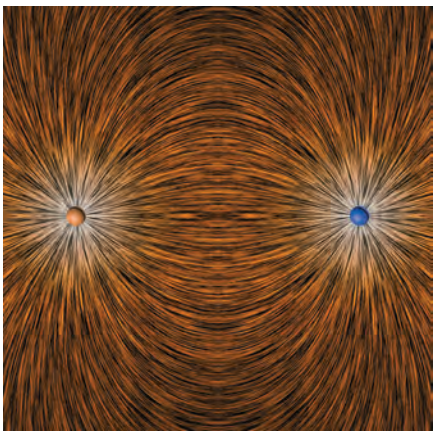


Figure 24.3 Two representations of the electric field of an electric dipole.

(*a*) Vector field diagram of an electric dipole



(*b*) Pattern created by electric dipole in a suspension of plastic fibers that align with the electric field



field lines is always proportional to the magnitude of the charge carried by the object. If, for example, 16 field lines emanate from an object that carries a charge $+q$, then 32 lines emanate from an object that carries a charge $+2q$ and eight lines terminate on an object that carries a charge $-q/2$.

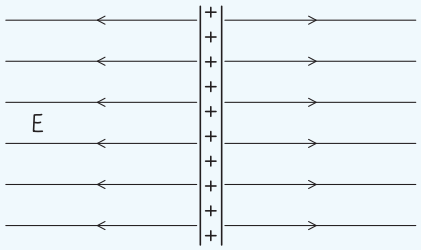
The number of field lines that emanate from a positively charged object or terminate on a negatively charged object is proportional to the charge carried by the object.

EXERCISE 24.1 Field lines of infinite charged plate

Draw a field line diagram for an infinite plate that carries a uniform positive charge distribution inside the plate.

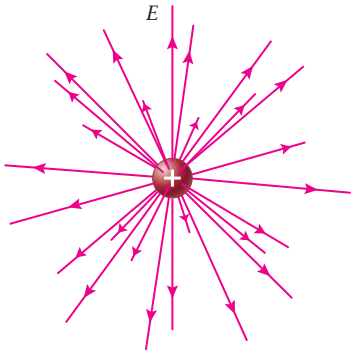
SOLUTION I know from Chapter 23 that the electric field produced by a charged plate of infinite area is always perpendicular to the plate (see Figure 23.30). Thus, the field lines must be straight lines perpendicular to the plate. Because the plate is positively charged, I draw the field lines perpendicular to and away from the plate on either side (**Figure 24.4**). ✔

Figure 24.4



Like vector field diagrams, field line diagrams provide an incomplete view of the electric field and are awkward to draw for all but the simplest charge distributions. Both types of diagram are limited by the two-dimensional nature of illustrations. In particular, you should keep in mind that field lines emanate in three dimensions (**Figure 24.5**), not just in the plane of the drawing.

Figure 24.5 Although we generally use two-dimensional representations of field line diagrams, field lines emanate in three dimensions.



CONCEPTS

24.2 (*a*) Is it possible for two electric field lines to cross? (Hint: What is the direction of the electric field at the point of intersection?) (*b*) Can two electric field lines touch?

24.2 Field line density

The most remarkable feature of field lines is this: Even though we take into account only the *direction* of the electric field when drawing field lines, they also contain information about the *magnitude* of the electric field. Figure 24.5 shows that as the distance from the charged object increases, the field lines are spaced farther apart from one another. To see whether there is a quantitative correspondence between field line spacing and electric field magnitude, complete the next checkpoint.

24.3 Imagine a hollow sphere enclosing the charged object in Figure 24.5, centered on the object. (*a*) Given that 26 field lines emanate from the charged object, how many field lines cross the surface of the hollow sphere? (*b*) If the radius of the hollow sphere is R, what is the number of field line crossings per unit surface area? (*c*) Now consider a second sphere with radius $2R$, also centered on the charged object. How many field lines cross the surface of this second sphere? (*d*) How does the number of field line crossings per unit area on the second sphere compare with that on the first sphere? (*e*) How does the electric field at a location on the second sphere compare with the field at a location on the first sphere?

From Checkpoint 24.3, we see that the electric field and the number of field line crossings per unit area both decrease as $1/r^2$. To express this correspondence quantitatively, we define a new quantity, the **field line density:**

The field line density at a given position is the number of field lines per unit area that cross a surface perpendicular to the field lines at that position.

Figure 24.6 illustrates why the surface through which the field lines pass must be perpendicular to the field lines. The field represented by the field lines in the figure is uniform—its magnitude and direction are the same everywhere. As you can see in the figure, the number of field lines that cross the surface depends on the orientation of the surface. The number of field lines that cross the surface is maximum when the surface is perpendicular to the field lines and decreases for any other orientation. We shall see later in this chapter how to account for the orientation of a surface when calculating field line density.

Because the number of field lines in a field line diagram is arbitrary, the field line density is also an arbitrary number, and so you may be wondering why field line density is a useful quantity. As you will see shortly, however, the field line density allows us to draw conclusions about electric field magnitudes. The only condition we make is that, in a given field line diagram, the number of field lines emanating from or terminating on charged objects is proportional to the magnitude of the charge carried by these objects.

24.4 (*a*) In Figure 24.6, for what orientation is the number of field lines that cross the surface a minimum? (*b*) How many field lines cross a plane surface of area 0.5 m^2 placed perpendicular to the field lines in Figure 24.6? (*c*) Using your answer to part *b*, what is the number of field line crossings *per unit area* through the 0.5-m^2 surface? (*d*) How does this compare to the number of field line crossings per unit area for the 1-m^2 surface in Figure 24.6*a*?

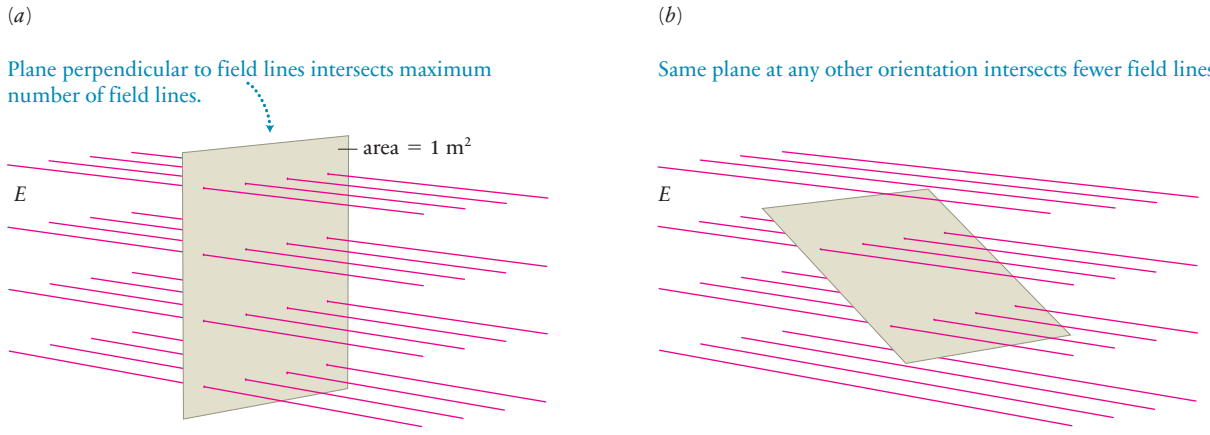
For the spherical surfaces of Checkpoint 24.3, the field lines are all perpendicular to the surface because the field lines are radial. The number of field line crossings you calculated per unit area *is* the field line density. These results lead us to conclude:

At every position in a field line diagram, the magnitude of the electric field is proportional to the field line density at that position.

The box "Properties of electric field lines" summarizes the properties of electric field lines.

CONCEPTS

Figure 24.6 The number of field lines that cross a given surface depends on the orientation of the surface relative to the field lines.



Properties of electric field lines

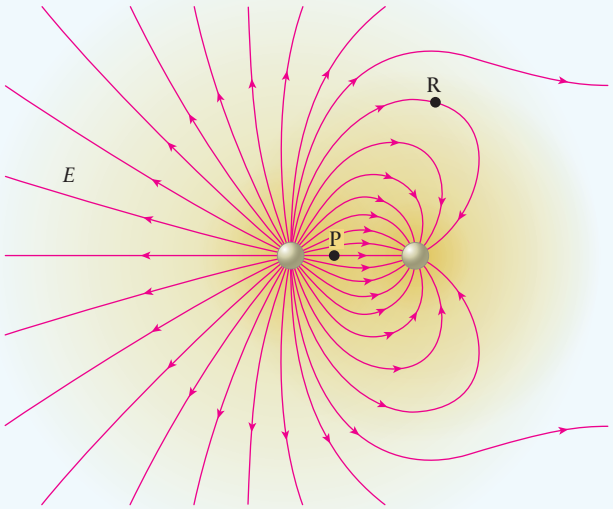
When working with electric field lines, keep the following points in mind:

1. Field lines emanate from positively charged objects and terminate on negatively charged objects.
2. At every position, the direction of the electric field is given by the direction of the tangent to the electric field line through that position.
3. Field lines never intersect or touch.
4. The number of field lines emanating from or terminating on a charged object is proportional to the magnitude of the charge on the object.
5. At every position, the magnitude of the electric field is proportional to the field line density.

EXERCISE 24.2 Field strength from field lines

Consider the field line diagram shown in **Figure 24.7**. (*a*) What are the signs of the charges on the two small spherical objects? (*b*) What are the relative magnitudes of these charges? (*c*) What is the ratio of the magnitudes of the electric fields at points P and R? (*d*) Is the electric field zero anywhere in the region shown?

Figure 24.7 Exercise 24.2.



SOLUTION (*a*) Because field lines leave the left object and terminate on the right one, the left object carries a positive charge and the right object carries a negative charge. ✔

(*b*) From Figure 24.7 I see that about twice as many field lines leave the left object as end on the right one. Thus, the charge on the left object is about twice that on the right object.* ✔

(*c*) The magnitude of the electric field at each position is proportional to the field line density at that position. The field line density is equal to the number of field lines per unit length, which is proportional to the inverse of the distance between adjacent field lines. Measuring with a ruler, I see that the distances between adjacent field lines at points P and R are $d_P \approx 1$ mm and $d_R \approx 6$ mm, so

$$\frac{E_P}{E_R} = \frac{d_R}{d_P} \approx \frac{6\text{ mm}}{1\text{ mm}} = 6. \checkmark$$

(*d*) The absence of field lines on the right suggests that the electric field is small (or even zero). Indeed, if a test particle carrying a positive charge is placed in that region, it is subject to a repulsive force exerted by the positively charged object on the left and an attractive force exerted by the negatively charged object on the right. If the test particle is to the right of the particles and $\sqrt{2}$ as far from the positively charged particle as it is from the negatively charged particle, the vector sum of the two forces is zero and so the electric field at that position is zero. ✔

*If you answered that the magnitude of the charge on the left is four times that on the right to account for three dimensions instead of two, imagine rotating the plane of the cross-section shown in Figure 24.7 around an axis that contains both charges. That cross-section won't change. So, even in three dimensions, the charge magnitude on the left is about twice that on the right.

Note that in Exercise 24.2, half of the field lines leave the area of interest. These field lines either eventually terminate on a negatively charged object (not shown) or continue out to "infinity."

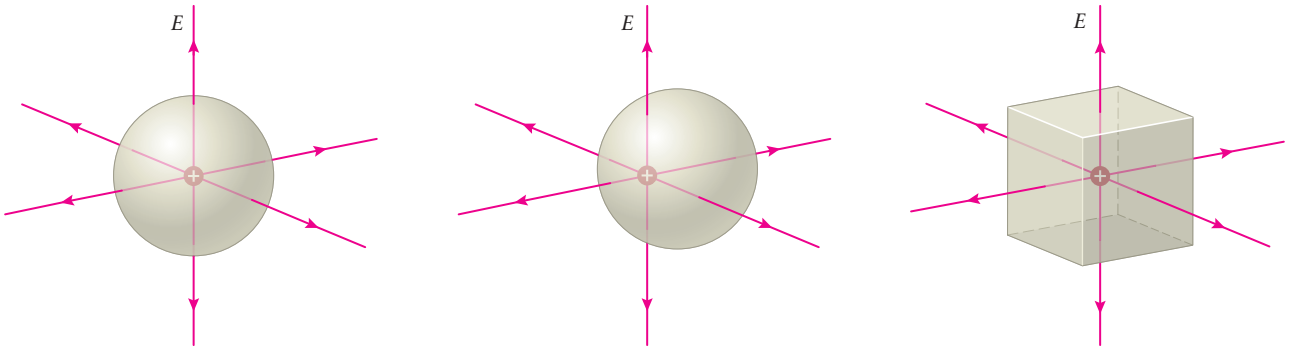
CONCEPTS

24.5 Imagine moving the hollow sphere of radius R of Checkpoint 24.3*a* sideways so that the charged object is no longer at the center of the sphere (but still within it). (*a*) How does the number of field line crossings through the surface of the sphere change as it is moved? (*b*) How does the average number of field line crossings per unit surface area of the sphere change? (*c*) Does the electric field strength at a fixed position on the surface of the sphere change or remain the same as the sphere is moved? (*d*) Are your answers to parts *b* and *c* in contradiction, given the relationship between the electric field magnitude and field line crossings per unit area?

24.3 Closed surfaces

Checkpoint 24.5 leads us to another result that will be important in deriving Gauss's law: Whenever a charged particle is placed inside a hollow spherical surface, the number of field lines that pierce the surface is the same *regardless of where inside the surface the particle is placed*. This is true simply because so long as the charged particle is inside the surface, all the field lines emanating from the particle must go

Figure 24.8 Any surface that encloses a positively charged particle is pierced by all the field lines that emanate from that particle, regardless of the shape of the surface and the position of the particle within the surface.



through the spherical surface. In fact, we don't even need to use a spherical surface—a cube-shaped surface or any other surface enclosing the charged particle will do (**Figure 24.8**). In each case, the number of field lines that pierce the surface is equal to the number of field lines that emanate from or terminate on the charged particle enclosed by the surface.

24.6 Suppose eight field lines emanate from an object carrying a charge $+q$. How many field lines pierce the surface of a hollow sphere if the sphere contains (*a*) a single object carrying a charge $+2q$ and (*b*) two separate objects, each carrying a charge $+q$? (*c*) If the sphere is pierced by 20 field lines, what can you deduce about the amount of combined charge on objects inside the sphere?

A surface that completely encloses a volume is called a **closed surface**. Checkpoint 24.6 suggests that a direct relationship exists between the number of field lines that cross a closed surface and the **enclosed charge**—the sum of all charge enclosed by that surface. However, what happens if a field line reenters the closed surface, as illustrated in **Figure 24.9a**? Field line 4 now crosses the closed surface *three* times, so the number of field line crossings is not six but eight. If you look closely at the figure, however, you will discover that not all the crossings are the same. For seven of the crossings the field line goes outward (from the inside of the closed surface to the outside), while for the eighth crossing the field line goes inward. If we assign a value of $+1$ to each outward crossing and a value of -1 to each inward crossing, we obtain $(+7) + (-1) = 6$.

To keep track of the number of inward and outward field crossings, we define a new quantity called the *field line flux*:

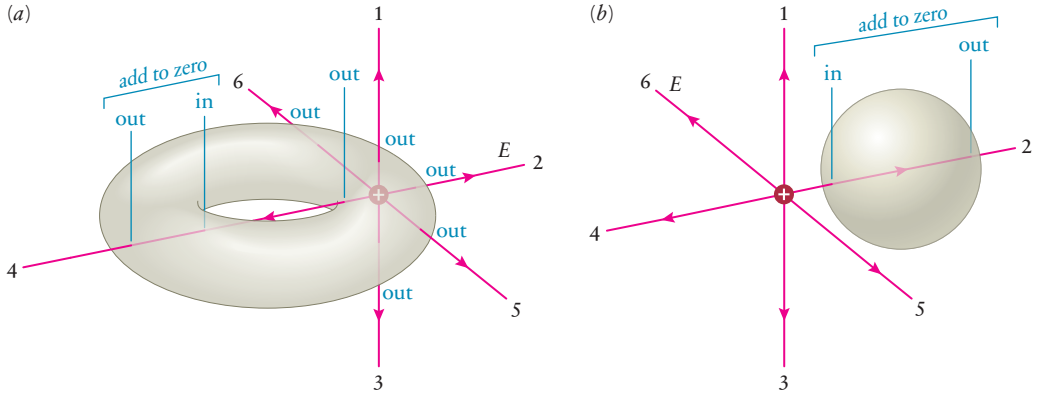
> **For any closed surface, the *field line flux* is the number of outward field lines crossing the surface minus the number of inward field lines crossing the surface.**

In calculating the field line flux for any closed surface, we assign a value of $+1$ to each outward field line crossing the surface and a value of -1 to each inward field line crossing the surface.

24.7 (*a*) If more than one field line reenters the donut in Figure 24.9*a*, what happens to the field line flux? (*b*) Are there any closed surfaces enclosing a charged particle through which the field line flux is different from that through a simple sphere around that particle?

CONCEPTS

Figure 24.9 The number of field lines exiting a closed surface minus the number entering it is always equal to the number of field lines generated inside the surface.



The field line flux through any closed surface is always equal to the number of field lines that originate from within that surface minus the number of field lines that terminate on charged objects within that surface. So far, we have drawn an arbitrary number of field lines, but we can make our statement more precise:

The *field line flux* through a closed surface is equal to the charge enclosed by the surface multiplied by the number of field lines per unit charge.

What about charged objects outside the closed surface? To see what effect such charged objects have, complete the following checkpoint.

24.8 (*a*) What is the field line flux through the closed spherical surface in Figure 24.9*b* due to a charged particle outside the sphere? (*b*) Does your answer to part *a* change if we move the particle around (but keep it outside the volume enclosed by the surface)?

Checkpoint 24.8 demonstrates a very important point:

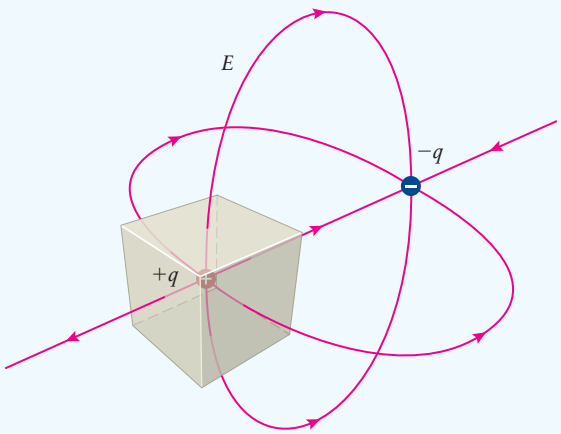
The field line flux through a closed surface due to charged objects outside the volume enclosed by that surface is always zero.

This means that if we know the field line flux through a closed surface, then we can determine the charge enclosed by that surface, regardless of the distribution of charge outside the surface. This statement is a form of Gauss's law, which we shall describe mathematically in Section 24.7.

EXAMPLE 24.3 Flux of an electric dipole

Consider the three-dimensional dipole field line diagram shown in **Figure 24.10**. Six field lines emanate from the positively charged end, and six terminate on the negatively charged end. (*a*) What is the field line flux through the surface of the cube that encloses the positively charged end shown in the figure? (*b*) What is the field line flux through the surface of a similar cube that encloses the negatively charged end?

Figure 24.10 Example 24.3.



SOLUTION (*a*) Six field lines emanate from the positively charged particle. Each line crosses the surface of the cube in the outward direction and thus contributes a value of +1 to the field line flux. The field line flux is +6. ✔

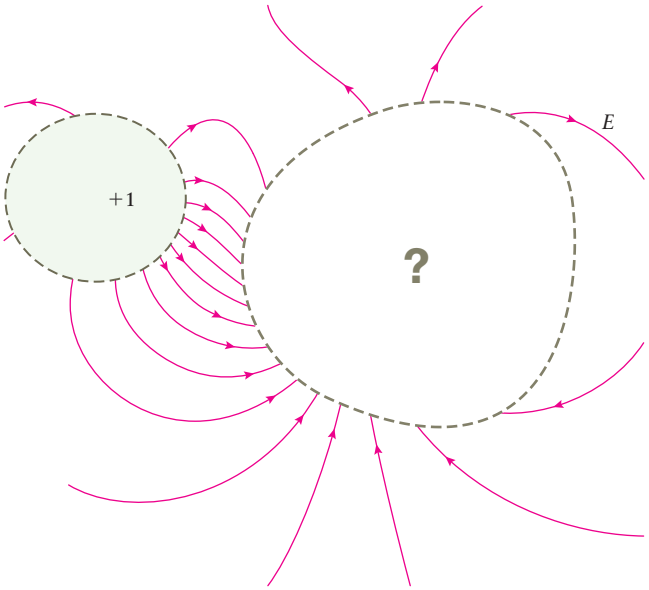
(*b*) Six field lines terminate on the negatively charged particle, so there are again six field line crossings. However, these field lines are directed inward and so the field line flux is −6. ✔

24.9 What is the field line flux through the surface of a rectangular box that encloses *both* ends of an electric dipole?

The relationship between the field line flux through a closed surface and the enclosed charge is important because it can help us determine one from a knowledge of the other. For example, in the next section we shall use this relationship to derive two important theorems about isolated conducting objects.

24.10 Consider the two-dimensional field line diagram in **Figure 24.11**, part of which is hidden from view. (*a*) If the object in the top left carries a charge of +1 C, what is the charge enclosed in the region that is hidden? (*b*) What is the field line flux through a surface that encloses the entire area represented by the diagram?

Figure 24.11 Checkpoint 24.10: What is the charge inside the dashed region?

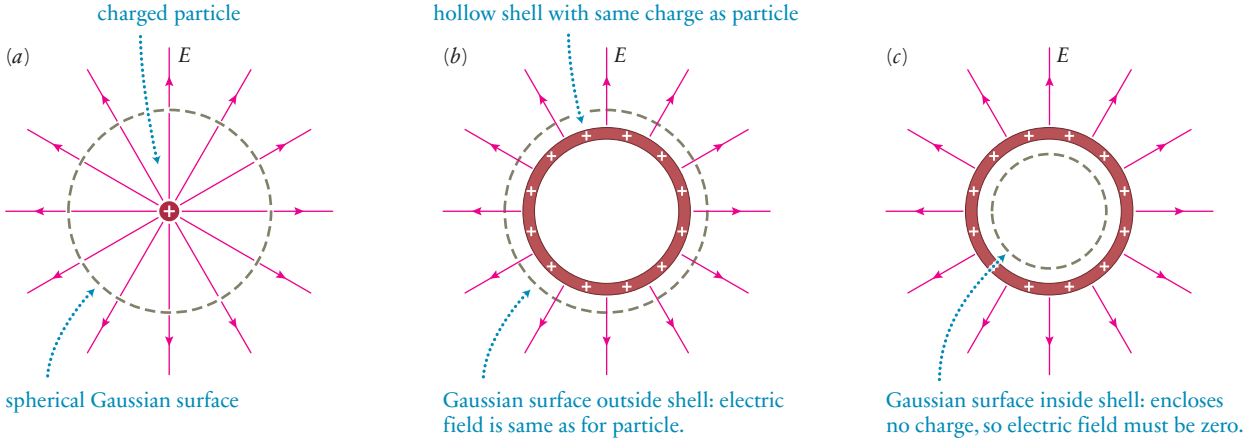


24.4 Symmetry and Gaussian surfaces

The relationship between field line flux and enclosed charge allows us to reach several important conclusions about charged objects and their electric fields without having to do any calculations. To apply this relationship

CONCEPTS

Figure 24.12 Using spherical Gaussian surfaces to examine the electric fields of a charged particle and a uniformly charged spherical shell. The electric fields, Gaussian surfaces, and charged shell are spherical and are shown here in cross section.



in a given situation, we first need to select a closed surface. This surface need not correspond to a real object—any surface, real or imagined, will do. We'll refer to these closed surfaces as **Gaussian surfaces.** The choice of surface is dictated by the symmetry of the situation at hand. As a rule of thumb, we choose a surface such that the electric field is the same (and possibly zero) everywhere along as many regions of the surface as possible, because such a choice makes it easy to determine the field line flux through the surface.

Consider, for example, the charged particle shown in **Figure 24.12*a***. As we have seen, the field lines for the particle radiate outward from it. (The figure shows only a two-dimensional cross section of the three-dimensional situation.) The field is symmetrical in all three dimensions—it has the same magnitude at the same distance from the center in any direction. Therefore, if we draw a spherical Gaussian surface that is concentric with the particle, the magnitude of the electric field is the same at all locations on the sphere. In other words, the field line density is the same all over the surface of the sphere. As we have seen in the preceding section, the field line flux through the Gaussian surface is proportional to the charge enclosed by the sphere.

Now suppose we replace the charged particle by a spherical shell that carries the same charge as the particle and still fits within our Gaussian surface (Figure 24.12*b*). If the charge is uniformly distributed over the shell, then the electric field should still be the same in all directions. Like the charged particle, the charged shell has *spherical symmetry* (see the box "Symmetry and Gauss's law"): Reorienting the spherical shell by rotating it over an angle about any axis does not change the charge configuration and so should not change the electric field at a given location. This means that the field lines should again be straight lines radiating uniformly outward. Also, the field line flux through the Gaussian surface should still be the same because the surface encloses the same amount of charge. The only way that the field line fluxes through the Gaussian surfaces in Figures 24.12*a* and *b* can be uniform *and* equal in magnitude is if the electric fields are the same at every position on the spherical Gaussian surface. Because this argument holds for a spherical Gaussian surface of any radius, we can conclude:

> **The electric field outside a uniformly charged spherical shell is the same as the electric field due to a particle that carries an equal charge located at the center of the shell.**

This means that a uniformly charged shell exerts a electric force on a charged particle outside the shell as if all the shell's charge were concentrated at the center of the shell. Because a sphere may be viewed as a collection of shells, we can extend this statement to uniformly charged spheres.

Let us now turn our attention to the electric field in the space enclosed by the shell. We draw a spherical Gaussian surface that fits within the shell (Figure 24.12*c*). This Gaussian surface encloses no charge, so the field line flux through the Gaussian surface is zero. Because the electric field can only be radially outward by symmetry, the electric field must be zero everywhere on the Gaussian surface. Because we can vary the radius of the Gaussian surface from zero to the inner radius of the shell without changing this argument, we can conclude:

> **In the absence of other charged objects, the electric field in the space enclosed by a uniformly charged spherical shell is zero everywhere in the enclosed space.**

Physically this means that a uniformly charged shell exerts no electric force on a charged particle located inside the shell.

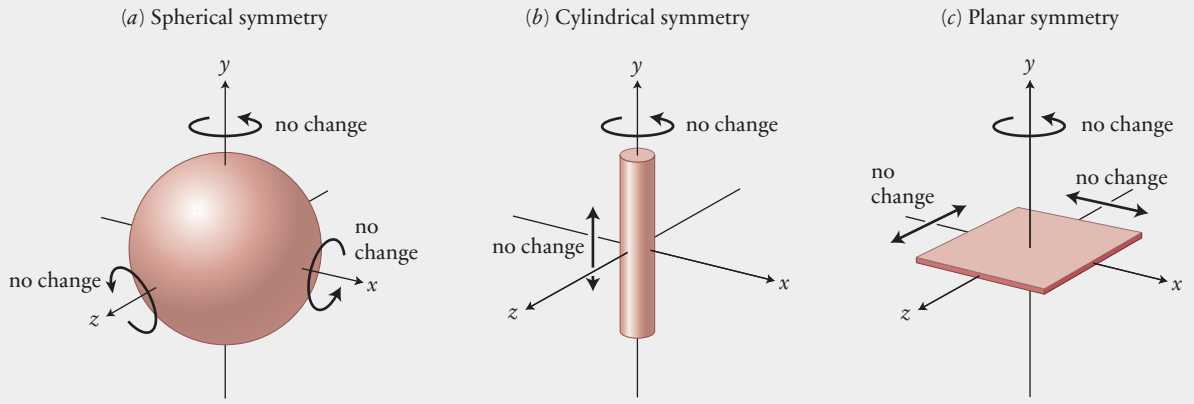
CONCEPTS

Symmetry and Gauss's law

The symmetry of an object is determined by its *symmetry operations*—manipulations that leave its appearance unchanged (see Section 1.2). A sphere, for example, looks the same if we reorient it by rotating it about any axis (Figure 24.13*a*). This type of symmetry is called **spherical symmetry.** An infinitely long, cylindrical rod does not look any different if we rotate it, reverse it, or translate it about its long axis (Figure 24.13*b*). The rod is said to have **cylindrical symmetry.** An infinite flat sheet has **planar symmetry:** It remains unchanged if it is rotated about an axis perpendicular to the sheet or translated along either of the two axes perpendicular to this axis (Figure 24.13*c*). Many other types of symmetry may occur, but these three types play an important role in electrostatics. For charge configurations that exhibit any of these three symmetries, we can calculate the electric field due to the charge distribution directly using Gauss's law.

Because objects are never infinite, they cannot exhibit true cylindrical or planar symmetry. However, for a long straight wire or a large flat sheet we can often obtain good results by assuming they have cylindrical or planar symmetry. When we work problems, the words *long* and *large* imply that you may assume the object has infinite dimensions compared to other length scales of interest.

Figure 24.13 Three symmetries important for applications of Gauss's law.

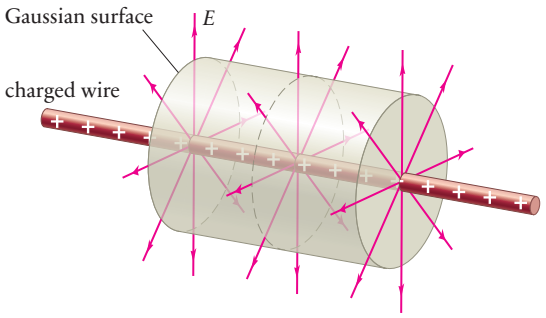


24.11 There are two reasons the field line flux through a closed surface may be zero: because the field is zero everywhere or because the outward flux is balanced by an equal inward flux. Why can't the latter situation be true for the Gaussian surface in Figure 24.12*c*?

Particles, shells, and spheres are the only objects that exhibit spherical symmetry. (Notice that unless specified otherwise, *shell* means *spherical shell*.) Figure 24.14 illustrates a different type of symmetry: the *cylindrical symmetry* of an infinitely long, uniformly charged straight wire.* Because of this symmetry, rotating the wire about its axis or moving it along the axis should not have any effect on the electric field at any position in space. For this to be the case, the field lines must be arranged radially along planes that are perpendicular to the wire (Figure 24.14). We can take advantage of this symmetry by drawing a Gaussian surface in the shape of a cylinder that is concentric with the wire, as shown in Figure 24.14.

*For the wire to exhibit cylindrical symmetry, it has to be infinitely long. If the wire has finite length, you can tell when it is moved along its axis.

Figure 24.14 The electric field of an infinite uniformly charged wire exhibits cylindrical symmetry. We can examine this field by surrounding the wire with a concentric cylindrical Gaussian surface.

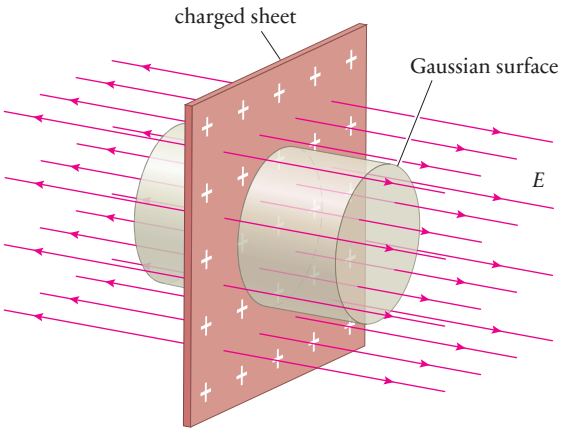


24.12 Consider a point on the curved part of the Gaussian surface in Figure 24.14. Does the magnitude of the electric field at that point increase, decrease, or stay the same if you (*a*) change the location of the point on the curved surface or (*b*) increase the radius of the Gaussian surface? (*c*) What is the field line flux through the left and right surfaces of the Gaussian surface?

We can use the cylindrical Gaussian surface to determine how the electric field due to the charged wire

CONCEPTS

Figure 24.15 The electric field of a uniformly charged sheet exhibits planar symmetry. We examine its field by drawing a cylindrical Gaussian surface that straddles the sheet.



decreases with distance from the wire. As you can see from the figure, the field line flux through the curved surface of the cylinder is independent of its radius r. No matter how large we make the radius of the cylinder, the same number of field lines pass through it. The area A of the curved surface is equal to the perimeter times the height h of the cylinder: $A = 2\pi rh$. As we increase the radius r of the cylinder, the surface area increases proportionally to r, and so the field line density must decrease as $1/r$ to maintain a constant number of field lines. Because the field line density is a measure of the electric field strength, this means that the electric field due to the wire decreases as $1/r$, as we established in Example 23.4. A quantitative expression for the electric field due to a charged rod is given in Section 24.8.

Figure 24.15 shows a situation with a different symmetry: a charged sheet. If the sheet is very large and the charge is uniformly distributed along it, then the electric field lines must be perpendicular to the sheet and also uniformly distributed along it. The one-dimensional symmetry exhibited by the electric field due to the charged sheet is an example of *planar symmetry*. To take advantage of this symmetry, we draw a cylindrical Gaussian surface that straddles the sheet, as illustrated in Figure 24.15.

CONCEPTS

24.13 Consider a point on the right surface of the Gaussian surface in Figure 24.15. Does the magnitude of the electric field at that point increase, decrease, or stay the same if you (*a*) change the location of the point on the right surface, or (*b*) increase the height h of the Gaussian surface? (*c*) Is the field line flux through the right surface of the Gaussian surface positive, negative, or zero? (*d*) How does the field line flux through the right surface compare to that through the left surface? (*e*) Is the field line flux through the curved surface of the Gaussian surface positive, negative, or zero?

Because the area of the right surface and the field line flux through that surface don't change as we change the height of the cylinder, we conclude that the electric field line density doesn't change with distance to the plane. Hence the magnitude of the electric field due to the charged sheet is the same everywhere (see also Example 23.6).

24.5 Charged conducting objects

Let as now apply the relationship between field line flux and enclosed charge to charged conducting objects. As we saw in Chapter 22, conducting materials permit the free flow of charge carriers within the bulk of the material. Conducting objects typically contain many charge carriers that are free to move, such as electrons (in a metal) or ions (in a liquid conductor). The material as a whole can still be electrically neutral; a neutral piece of metal, for example, contains as many positively charged protons as negatively charged electrons.

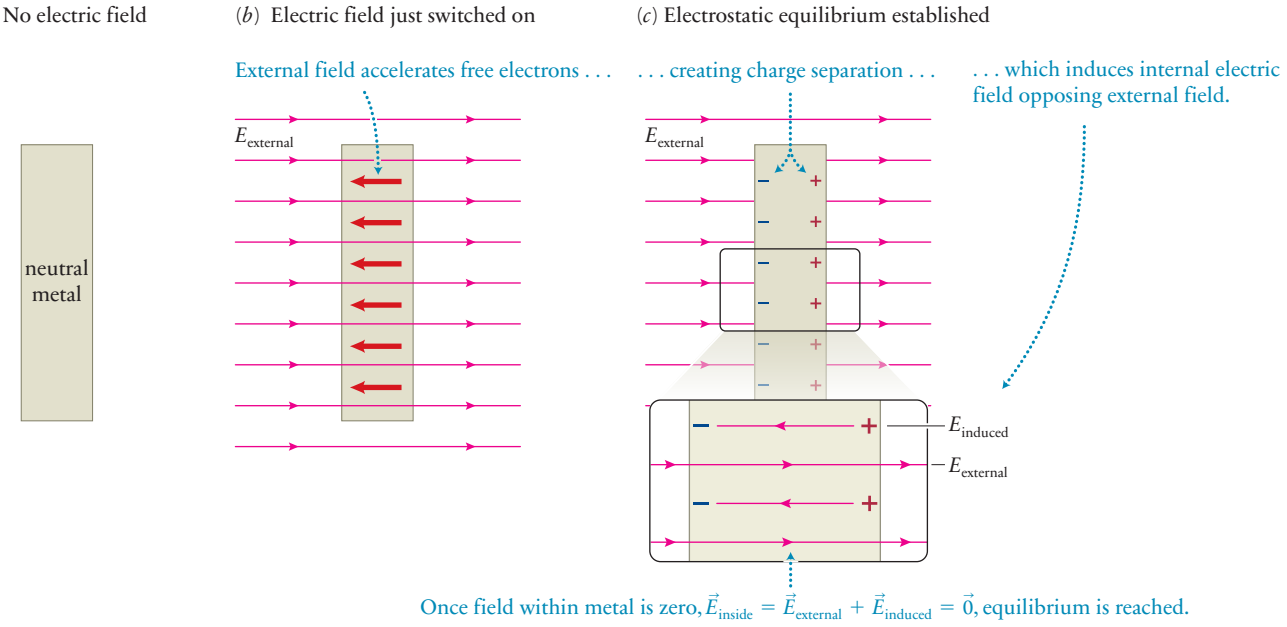
A consequence of this free motion of charged particles within a conducting object is that the particles always arrange themselves in such a way as to make the electric field inside the bulk of the object zero. To see how this comes about, consider a free electron in a slab of metal. If no field is present, no electric force is exerted on the electron. If we apply an external field, however, the free electron is subject to a force in a direction opposite the direction of the electric field (opposite because of the negative charge of the electron).

In a similar way, all the free electrons in a slab of metal initially accelerate in a direction opposite the direction of an applied field (**Figure 24.16**). This leaves behind a positive charge on one side of the slab and creates a negative charge on the opposite side. Because of this rearrangement of charge, an induced electric field builds up in a direction opposite the direction of the external field. As a result, the electric field inside the slab, which is the sum of the external electric field and the induced electric field, decreases. As this field decreases, so does the force exerted on the free electrons in the slab. When enough charge carriers have accumulated on each side of the slab to make the electric field inside the slab zero, the electric force exerted on the free electrons in the metal becomes zero and the material reaches **electrostatic equilibrium**—the condition in which the distribution of charge in a system does not change. The time interval it takes for a metal to reach electrostatic equilibrium is very short (about 10^{-16} s), so the rearrangement of charge carriers is virtually instantaneous. The important point to remember is:

The electric field inside a conducting object that is in electrostatic equilibrium is zero.

Keep in mind that this statement holds *only* in electrostatic equilibrium. When charge carriers are made to flow through a conducting object—as in any electric or electronic device, like your stereo or a refrigerator—the electric field is *not* zero inside the object!

Figure 24.16 Why the electric field inside the bulk of a conducting object is zero when the object is in electrostatic equilibrium.



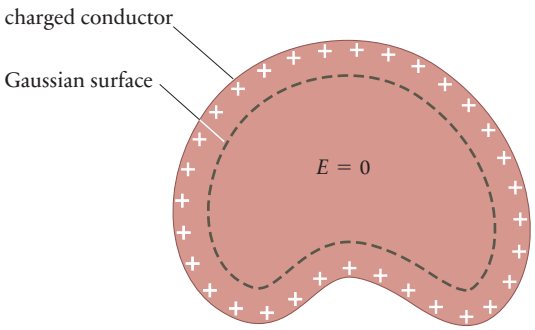
Suppose now we add charge to a conducting object. It makes sense to assume that the charged particles will arrange themselves over the object in such a way as to spread out as far as possible from one another, given that particles carrying like charges repel. We can use a Gaussian surface to obtain a better understanding of where the charged particles go.

24.14 Consider a spherical Gaussian surface inside a positively charged conducting object that has reached electrostatic equilibrium. (*a*) Is the field line flux through the Gaussian surface positive, negative, or zero? (*b*) What can you conclude from your answer to part *a* about the charge enclosed by the Gaussian surface?

We can extend the result of Checkpoint 24.14 to conducting objects of any shape. Consider, for example, the irregularly shaped, charged conducting object shown in **Figure 24.17**. Draw a Gaussian surface of the same shape as the object, just below its surface. Given that the field is zero everywhere inside the conducting object, the field line flux through the Gaussian surface is zero and the charge enclosed by the Gaussian surface is also zero. Because we can choose the Gaussian surface arbitrarily close to the surface of the object, we conclude:

Any surplus charge placed on an isolated conducting object arranges itself at the surface of the object. No surplus charge remains inside the conducting object once that object has reached electrostatic equilibrium.

Figure 24.17 Because the electric field inside a conducting object in electrostatic equilibrium is zero, we conclude that there cannot be any surplus charge inside the object.



24.15 Suppose the charged conducting object in Figure 24.17 contains an empty cavity. Does any surplus charge reside on the inner surface of the cavity?

CONCEPTS

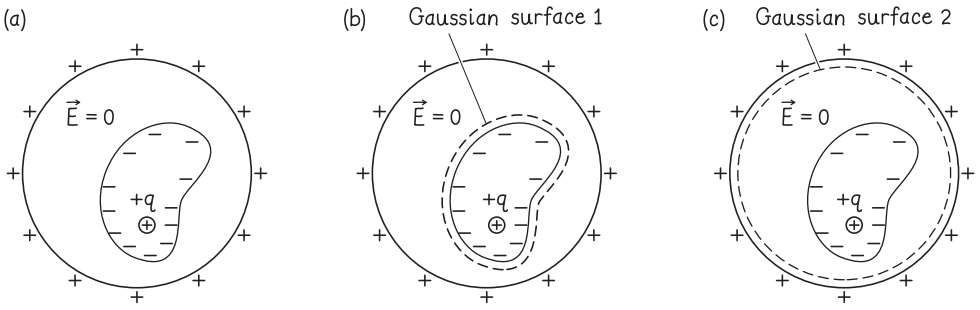
EXAMPLE 24.4 Charged particle in a cavity

An electrically neutral, conducting sphere contains an irregularly shaped cavity. Inside the cavity is a particle carrying a positive charge $+q$. What are the sign and magnitude of the charge on the sphere's outer surface?

1 GETTING STARTED I am told that a sphere made of material that is an electrical conductor has a cavity in its interior and that a particle in the cavity carries charge $+q$. My task is to determine the sign and magnitude of any charge residing on the sphere's outer surface. I begin by sketching a vertical cross section through the sphere showing the cavity and the

(Continued)

Figure 24.18



charged particle inside it (**Figure 24.18*a***). The problem states that the sphere is electrically neutral, but the question posed implies that some charge resides on its outer surface. Because $\vec{E} = \vec{0}$ inside the conducting material, I know that an equal quantity of the opposite charge must accumulate somewhere else on the sphere.

❷ **DEVISE PLAN** The sphere is conducting, so I know that once electrostatic equilibrium is reached, the electric field inside the bulk of the sphere must be zero: $\vec{E} = \vec{0}$. I can use this information to draw any Gaussian surface inside the sphere and use the following reasoning to determine the charge enclosed by my Gaussian surface: Because $\vec{E} = \vec{0}$ inside the bulk of the sphere and because I draw my Gaussian surface inside the sphere, $\vec{E} = \vec{0}$ everywhere on the Gaussian surface. Therefore the field line flux through the Gaussian surface is zero, which means the charge enclosed by this surface must be zero. Because I can draw my Gaussian surface anywhere inside the bulk of the conductor, I can use this information to determine the distribution of charge on the sphere.

❸ **EXECUTE PLAN** I begin by drawing a Gaussian surface 1 enclosing the cavity (Figure 24.18*b*). Because the field line flux through this surface is zero, the charge enclosed by the surface must be zero. There is charge $+q$ inside the surface (in the charged particle), however, and so in order for the charge enclosed by the surface to be zero, a quantity of charge $-q$ must have migrated from someplace in the region surrounding the cavity and accumulated on the inner cavity surface.

Because the sphere is electrically neutral, the charge $-q$ that migrated to the inner cavity surface must leave a charge $+q$ behind somewhere else on the sphere. If I now draw Gaussian surface 2 just inside the sphere's outer surface (Figure 24.18*c*), I see that, because the field line flux through this Gaussian surface has to be zero, the charge enclosed by this surface is also zero. The positive charge $+q$ that results from the migration of charge $-q$ to the cavity surface must therefore reside outside Gaussian surface 2. Because I can draw surface 2 arbitrarily close to the sphere's outer surface, I conclude that the sphere's outer surface carries a charge $+q$. ✔

❹ **EVALUATE RESULT** The negative charge that migrates from the region outside the cavity to the cavity surface arranges itself in such a way as to cancel the electric field that the charged particle creates in the region outside the cavity. Therefore all the field lines that start on the charged particle must end on the negative charge at the cavity surface. In order for all the field lines to end here, the quantity of negative charge on the cavity surface must be equal to the quantity of charge on the particle. The sphere is electrically neutral, and my choice for where I draw Gaussian surface 2 requires that all the positive charge resulting from the migration of negative charge to the cavity surface must accumulate outside Gaussian surface 2, which means right at the sphere's outer surface. Thus my answer makes sense.

That the electric field inside any conductor is zero in electrostatic equilibrium allows us to draw one additional important conclusion. Because the electric fields must be zero everywhere, including at the surface of a conducting object, there cannot be any component of the electric field parallel to the surface of the object, and therefore we can conclude:

In electrostatic equilibrium, the electric field at the surface of a conducting object is perpendicular to that surface.

If there were a component of the electric field parallel to the surface, that component would cause any free charge carrier to move along the surface, which means the conductor is not in electrostatic equilibrium.

CONCEPTS

24.16 In Example 24.4, is the electric field inside the cavity zero?

Self-quiz

1. In **Figure 24.19**, which of the two charged spheres carries a charge of greater magnitude?
2. Consider Gaussian surfaces 1–3 in Figure 24.19. Determine the field line flux through each surface (considering just the two dimensions shown).
3. In Figure 24.19, is the field line density greater at point A or point B? At which of these locations is the magnitude of the electric field greater? Is the field line density at point C zero or nonzero?
4. The electric field lines in **Figure 24.20** tell you there must be one or more charged particles inside the Gaussian surface defined by the dashed line. Could the electric field shown be due to a single particle inside the Gaussian surface? What must the signs and relative magnitudes of the charged particle(s) be in order to create the electric field lines shown?
5. **Figure 24.21** shows a small ball that carries a charge of $+q$ inside a conducting metal shell that carries a charge of $+2q$. (*a*) What are the sign and magnitude of the charge on the inner surface of the shell? (*b*) What are the sign and magnitude of the charge on the outer surface of the shell?

Figure 24.19

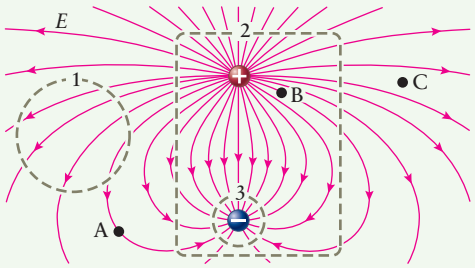


Figure 24.20

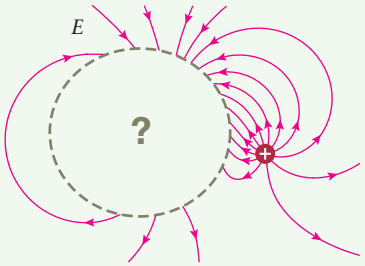
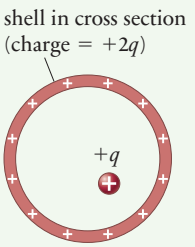


Figure 24.21



Answers

1. The number of field lines is proportional to the charge on the object. Because more lines emanate from the charged object on the top, that object must carry a greater charge.
2. For surface 1, all lines that enter the surface also exit the surface, so the field line flux is zero. For surface 2, 25 lines exit the surface and 6 lines enter the surface. The field line flux through surface 2 is thus $25 - 6 = 19$. Fifteen field lines cross surface 3, with all lines entering from the outside. The field line flux for surface 3 is -15.
3. Point B has the greater field line density because the lines are closer together at B than they are at A. The magnitude of the electric field is greater at point B because electric field strength is proportional to the field line density. Even though C is not on a field line, the field line density, which is represented by the spacing of the field lines *around* point C, is nonzero.
4. Because electric field lines converge on one point near the top of the area enclosed by the surface and diverge from a point near the bottom, there must be objects that carry both negative and positive charges inside the surface. Because more field lines enter the surface than exit the surface, the negatively charged object(s) must carry a charge of greater magnitude than the positively charged object(s).
5. (*a*) Because the electric field in the conducting shell is zero, the field line flux through a Gaussian surface drawn inside the material of the conducting shell must be zero. According to Gauss's law, the charge enclosed in the surface must also be zero. The charge on the inner surface of the shell must therefore be $-q$, which added to $+q$ gives zero. (*b*) For a neutral shell, a charge of $-q$ on the inside surface of the shell would leave a surplus of $+q$ on the outside surface of the shell. The surplus charge of $+2q$ that was placed on the shell also resides on the outside surface. The charge residing on the outside surface of the shell is thus $+3q$.

24.6 Electric flux

We introduced two important concepts in the first part of this chapter: the field line density, which is proportional to the strength of the electric field, and the field line flux, which represents the number of field lines going outward through a closed surface minus the number of field lines going inward. In this section we'll turn these concepts into quantities we can calculate.

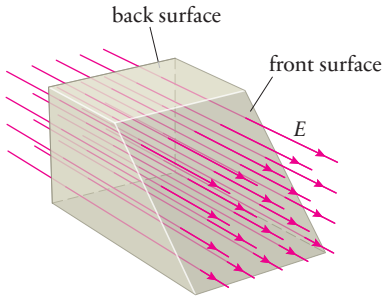
Consider, for example, a trapezoidal box in a uniform electric field $\vec{E}$ (**Figure 24.22***a*). The field line flux through the closed surface of the trapezoidal box is zero: Twenty field lines go into the back surface and 20 come out through the front surface. Another way of putting this is to say that the field line flux into the back surface is equal in magnitude to the field line flux out of the front surface. Instead of using field lines, however, whose number is chosen arbitrarily, we'll work with a quantity called the **electric flux,** represented by the symbol Φ_E (Φ is the Greek capital phi). The magnitude of the electric flux through a surface with area A in a uniform electric field of magnitude E is defined as

$$\Phi_E \equiv EA\cos\theta \quad \text{(uniform electric field)}, \tag{24.1}$$

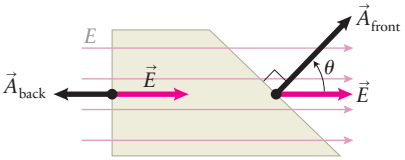
where θ is the angle between the electric field and the normal to the surface.

Figure 24.22 Determining the electric flux through a flat surface.

(*a*) Trapezoidal box in uniform electric field



(*b*) Vector areas of front and back surfaces



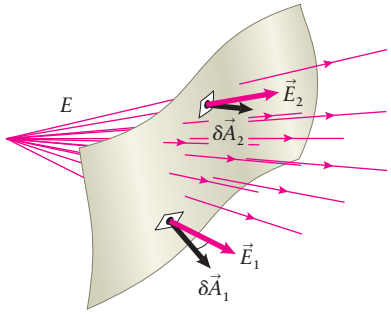
24.17 (*a*) Consider the front surface of the trapezoidal box in Figure 24.22*a*, detached from the rest of the trapezoidal box. Does the field line flux through that surface increase, decrease, or stay the same if any of the following quantities is increased: (*i*) the area of the front surface, (*ii*) the magnitude of the electric field (keeping the area constant), (*iii*) the slope of the front surface (that is to say, the angle between the surface and the direction of the electric field is increased)? (*b*) Does the *electric flux* through the front surface increase, decrease, or stay the same if any of these quantities is changed?

As Checkpoint 24.17 shows, electric flux, as we've defined it in Eq. 24.1, behaves like the field line flux. To make the correspondence more precise, we define an *area vector* $\vec{A}$ for a flat surface area as a vector whose magnitude is equal to the surface area A and whose direction is normal to the plane of the area. On closed surfaces we choose $\vec{A}$ to point outward (we'll deal with open surfaces later). Figure 24.22*b* shows the area vectors associated with the front and back surfaces of the (closed) trapezoidal box. With this definition, Eq. 24.1 can be written as a scalar product (Eq. 10.33):

$$\Phi_E \equiv EA\cos\theta = \vec{E}\cdot\vec{A} \quad \text{(uniform electric field)}, \tag{24.2}$$

where θ is the angle between $\vec{E}$ and $\vec{A}$. Electric flux is a scalar, and the SI unit of electric flux is $\text{N}\cdot\text{m}^2/\text{C}$.

Figure 24.23 To obtain the electric flux through an irregularly shaped, nonplanar surface and/or for a nonuniform electric field, we divide the surface into small segments. For very small segments, each segment is essentially flat and the field through each segment is essentially uniform. The flux through the entire surface is then given by the sum of all of the contributions through each segment.



24.18 Let the area of the back surface of the trapezoidal box in Figure 24.22 be $1.0\ \text{m}^2$, the magnitude of the electric field be $1.0\ \text{N/C}$, and $\theta = 30°$ for the front surface. (*a*) What are the magnitudes of the area vectors for the front and back surfaces of the trapezoidal box? (*b*) What are the electric fluxes through the front and back surfaces?

The above definition of electric flux applies only for uniform electric fields and flat surfaces. Let us therefore consider the more general case of an irregular surface in a nonuniform field (**Figure 24.23**). To calculate the electric flux through that surface, we divide the entire surface into small segments of surface area δA_i, with each segment being small enough that we

QUANTITATIVE TOOLS

can consider it to be essentially flat, and we can define an area vector $\delta\vec{A}_i$ whose magnitude is equal to the surface area $\delta\vec{A}_i$ of the segment and whose direction is normal to the segment. This allows us to apply Eq. 24.2 to each individual segment. The electric flux through a single segment is then $\Phi_{Ei} = \vec{E}_i \cdot \delta\vec{A}_i$, where $\vec{E}_i$ is the electric field vector at the location of the segment. To calculate the electric flux through the entire surface we must sum the electric flux through all the surface segments:

$$\Phi_E = \Sigma \vec{E}_i \cdot \delta\vec{A}_i. \tag{24.3}$$

If we let the area of each segment approach zero, then the number of segments approaches infinity and the sum is replaced by an integral:

$$\Phi_E = \lim_{\delta A_i \to 0} \Sigma \vec{E}_i \cdot \delta\vec{A}_i = \int \vec{E} \cdot d\vec{A}. \tag{24.4}$$

The integral in Eq. 24.4 is called a *surface integral* (see Appendix B); $d\vec{A}$ is the area vector of an infinitesimally small surface segment. If the surface is closed, this surface integral is written as

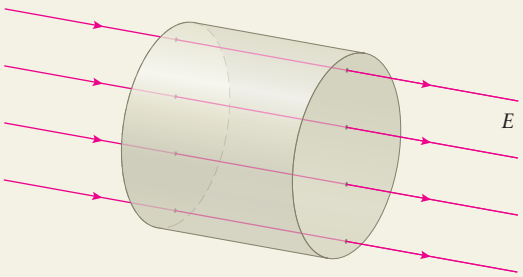
$$\Phi_E = \oint \vec{E} \cdot d\vec{A}, \tag{24.5}$$

where the circle through the integral sign indicates that the integration is to be taken over the entire closed surface and $d\vec{A}$ is chosen to point outward. Because evaluating a surface integral is mathematically more complicated than single-variable integration, it is important to exploit any symmetry to simplify the calculation.

EXAMPLE 24.5 Cylindrical Gaussian surface in a uniform electric field

Consider a cylindrical Gaussian surface of radius r and length ℓ in a uniform electric field $\vec{E}$, with the length axis of the cylinder parallel to the electric field (**Figure 24.24**). What is the electric flux Φ_E through this Gaussian surface?

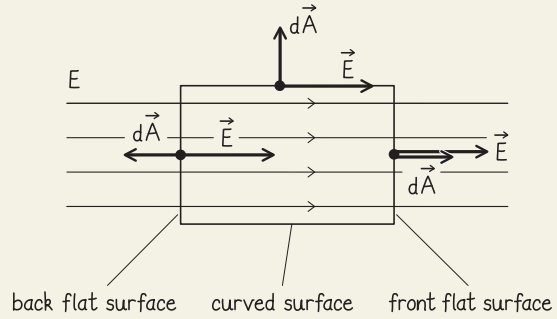
Figure 24.24 Example 24.5.



❶ **GETTING STARTED** From Figure 24.24 I see that a cylindrical Gaussian surface consists of three regions: front and back flat surfaces and a curved surface joining them. The electric field is perpendicular to the front and back flat surfaces and parallel to the curved surface.

❷ **DEVISE PLAN** The electric flux is given by Eq. 24.5, so I can calculate the electric flux through the Gaussian surface by applying Eq. 24.5 to each of the three regions and then summing the three contributions to the electric flux. In order to evaluate the scalar product $\vec{E} \cdot d\vec{A}$ for each region, I sketch a side view of the Gaussian surface showing the vectors $\vec{E}$ and $d\vec{A}$ for each of the three regions (**Figure 24.25**).

Figure 24.25



❸ **EXECUTE PLAN** Applying Eq. 24.5 to the three regions, I can write the flux through the Gaussian surface as the sum of three surface integrals: one over the back surface, one over the curved surface, and one over the front surface:

$$\Phi_E = \oint \vec{E} \cdot d\vec{A} = \underbrace{\int \vec{E} \cdot d\vec{A}}_{\text{back flat surface}} + \underbrace{\int \vec{E} \cdot d\vec{A}}_{\text{curved surface}} + \underbrace{\int \vec{E} \cdot d\vec{A}}_{\text{front flat surface}}. \tag{1}$$

From my sketch, I see that on the back flat surface the angle between $\vec{E}$ and $d\vec{A}$ is 180°, so $\vec{E} \cdot d\vec{A} = E(\cos 180°)\, dA = -E\, dA$. Because the magnitude of the electric field is the same everywhere, I can pull E out of the integral:

$$\underbrace{\int \vec{E} \cdot d\vec{A}}_{\text{back flat surface}} = \int (-E)\, dA = -E \int dA = -E(\pi r^2),$$

where πr^2 is the area of the back flat surface.

(Continued)

QUANTITATIVE TOOLS

The integral over the curved region of the Gaussian surface yields a value of zero because the angle between between $\vec{E}$ and $d\vec{A}$ is 90° everywhere on the curved region, so $\vec{E} \cdot d\vec{A} = E(\cos 90°)\, dA = 0$.

Finally, for the front flat surface I have $\vec{E} \cdot d\vec{A} = E(\cos 0°)\, dA = E\, dA$, so

$$\int_{\substack{\text{front flat}\\ \text{surface}}} \vec{E} \cdot d\vec{A} = \int E\, dA = E \int dA = E(\pi r^2).$$

Adding up the three terms in Eq. 1 yields

$$\Phi_E = -E(\pi r^2) + 0 + E(\pi r^2) = 0. \checkmark$$

4 EVALUATE RESULT Because there is no charge enclosed by the Gaussian surface, I know that the field line flux through the surface must be zero, a fact I confirm by looking at Figure 24.24: The four field lines shown contribute a flux of −4 on the back flat surface, +4 on the front flat surface, and 0 along the curved surface, for a total of $-4 + 4 + 0 = 0$. It therefore makes sense that the electric flux through this Gaussian surface also is zero.

24.19 Consider a spherical Gaussian surface of radius r with a particle that carries a charge $+q$ at its center. (*a*) On the Gaussian surface, what is the magnitude of the electric field due to the particle? (*b*) What is the electric flux through the sphere due to the charged particle? (*c*) Combining your answers to parts *a* and *b*, what is the relationship between the electric flux through the sphere and the enclosed charge q_{enc}? (*d*) Would this relationship change if you doubled the radius r of the sphere?

24.7 Deriving Gauss's Law

Checkpoint 24.19 shows that the electric flux through a spherical Gaussian surface is equal to the charge q enclosed by the sphere times $4\pi k$, where $k = 9.0 \times 10^9\ \text{N} \cdot \text{m}^2/\text{C}^2$ is the proportionality constant that appears in Coulomb's law (see Eqs. 22.1 and 22.5). This relationship is usually written in the form

$$\Phi_E = 4\pi k q = \frac{q}{\epsilon_0}, \tag{24.6}$$

where ϵ_0 is called the **electric constant:**

$$\epsilon_0 \equiv \frac{1}{4\pi k} = 8.85418781 \times 10^{-12}\ \text{C}^2/(\text{N} \cdot \text{m}^2). \tag{24.7}$$

Equation 24.6 is a special case of **Gauss's law,** which states that the electric flux through the closed surface of an arbitrary volume is

$$\Phi_E = \oint \vec{E} \cdot d\vec{A} = \frac{q_{\text{enc}}}{\epsilon_0}, \tag{24.8}$$

where q_{enc} is the **enclosed charge**—the sum of all charge on an object or portion of an object enclosed by the closed surface. The formal proof of Gauss's law is an extension of the calculation you performed in Checkpoint 24.19 and is shown in the box "Electric flux though an arbitrary closed surface."

Gauss's law is a direct consequence of Coulomb's law with its $1/r^2$ dependence and the superposition of electric fields. In that respect, it contains nothing new. However, as the next section shows, Gauss's law greatly simplifies the calculation of electric fields due to charge distributions that exhibit one of the three symmetries we discussed in Section 24.4.

24.20 Suppose Coulomb's law showed a $1/r^{2.00001}$ dependence instead of a $1/r^2$ dependence. (*a*) Calculate the electric flux through a spherical Gaussian surface of radius R centered on a particle carrying a charge $+q$. (*b*) Substitute your result in Eq. 24.8. What do you notice?

QUANTITATIVE TOOLS

Electric flux through an arbitrary closed surface

Consider a particle that carries a positive charge $+q$, surrounded by the irregularly shaped closed surface shown in **Figure 24.26*a***.

1. To determine the electric flux through the irregular surface, we divide the volume enclosed by the surface into small square wedges that taper to a point at the charged particle, one of which is shown. We calculate the electric flux through the surface segment dA cut out by each wedge and then sum the contributions from all the wedges.
2. To determine the electric flux through dA in Figure 24.26*a*, we draw two spherical Gaussian surfaces around q: one with a radius r_1 equal to the distance between dA and q, the other with an arbitrary radius r_2. Our wedge from step 1 now defines two other small surface segments dA_1 and dA_2 on these two spheres.
3. If the segment dA is made very small, then the field lines in the wedge are nearly parallel to one another. Therefore, according to what we found in Section 24.6, the electric flux through dA is equal to that through dA_1 (Figure 24.26*b*). In addition, as you showed in Checkpoint 24.19, this flux is also equal to that through surface segment dA_2. In fact, the electric flux is the same through *any* surface that cuts through the wedge. Put differently, any surface that cuts through the wedge intercepts the same number of field lines.
4. We can repeat this procedure for each wedge. Each time, we see that the electric flux through a segment dA on the irregular surface is equal to that through a corresponding segment dA_2. As we add the contributions of all the wedges, we conclude that the electric flux through the closed irregular surface is equal to that through a sphere of arbitrary radius r centered on q.
5. If there is more than one charged particle inside the irregularly shaped surface, we can use the above arguments for each particle individually and then use the superposition of electric fields (Section 23.3):

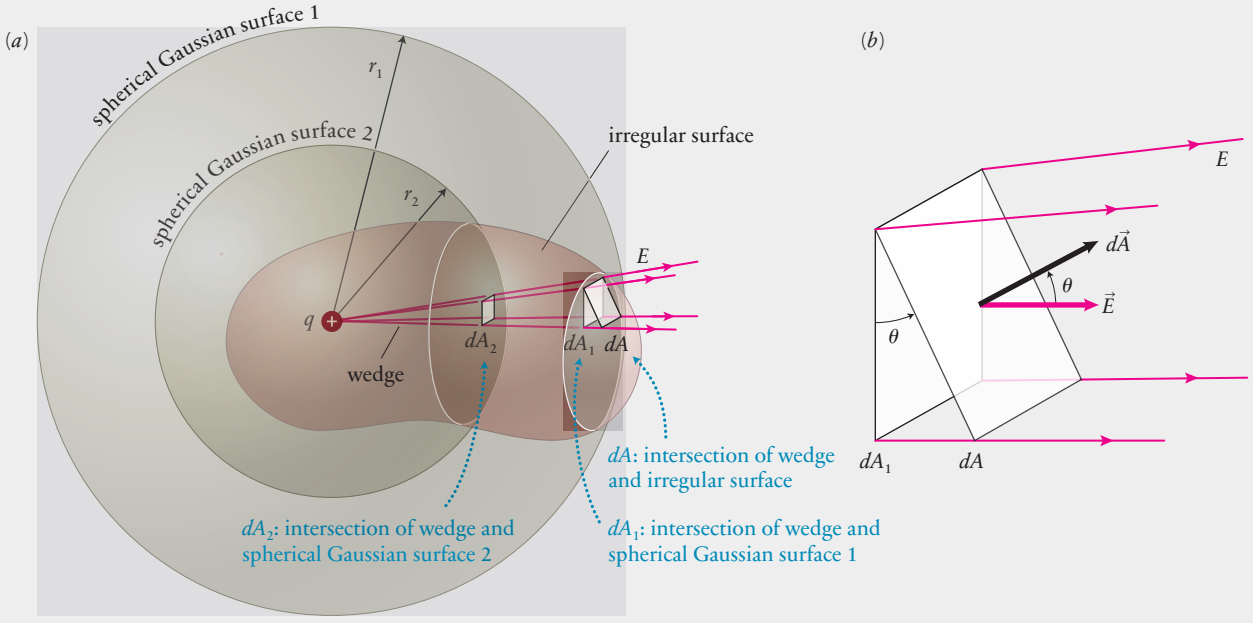
$$\vec{E} = \Sigma \vec{E}_i,$$

where $\vec{E}_i$ is the electric field due to particle i alone. The electric flux due to all charged particles is then the sum of the electric fluxes due to the individual electric fields:

$$\Phi_E = \Sigma \Phi_{Ei} = \Sigma \frac{q_i}{\epsilon_0} = \frac{q_{\text{enc}}}{\epsilon_0}.$$

24.21 Suppose q is *outside* the irregularly shaped surface in Figure 24.26. Show that the electric flux due to q through the closed surface is zero. (Hint: Draw a small wedge from q through the surface and determine the electric flux through the two intersections between the wedge and the surface.)

Figure 24.26 Formal proof of Gauss's law.



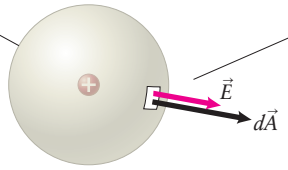
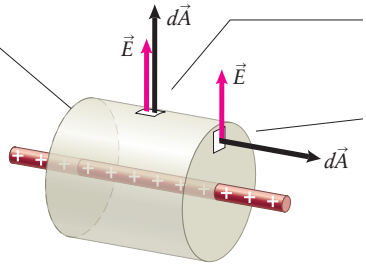
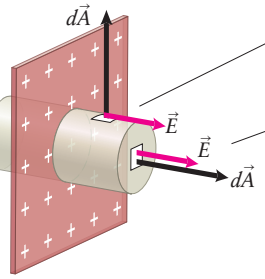
24.8 Applying Gauss's Law

Gauss's law relates the electric flux through a closed surface to the charge enclosed by it. In Section 24.5 we encountered one important application of Gauss's law: If we can choose a Gaussian surface such that the electric field is zero everywhere on that surface (inside a conducting material, for example), then we know that the charge enclosed by that surface must be zero. In this section we show how Gauss's law can be used to avoid having to carry out any integrations to calculate the electric field. In principle one can calculate the electric flux through any surface, but the calculation is not trivial in general. For the surfaces and charged objects shown in Section 24.4 and summarized in **Figure 24.27**, however, the electric flux is easy to calculate because the field lines are either parallel to the surface (in which case the electric flux is zero) or perpendicular to it and the magnitude of the electric field is constant (in which case the electric flux is simply the product of the magnitude of the electric field E and the surface area A).

To see the benefit of Gauss's law, consider a charged spherical shell of radius R that carries a uniformly distributed positive charge q. The electric field due to this charged shell can be calculated using the procedure outlined in Section 23.7: Divide the shell into infinitesimally small segments that carry a charge dq, apply Coulomb's law to each small segment, and integrate over the entire shell (Eq. 23.15):

$$\vec{E}_{\text{sP}} = k \int \frac{dq_{\text{s}}}{r_{\text{sP}}^2} \hat{r}_{\text{sP}}. \tag{24.9}$$

Figure 24.27 Applying Gauss's law to determine the electric fields of symmetrical charge distributions.

Symmetry of charge distribution	Electric field geometry	Gaussian surface	To find electric flux
Spherical (charged sphere)	$\vec{E}$ radiates uniformly outward in three dimensions.	Concentric sphere 	At all points, $\vec{E}$ is perpendicular to surface and has same magnitude.
Cylindrical (infinite charged rod)	$\vec{E}$ radiates uniformly outward perpendicular to axis.	Coaxial cylinder 	*Cylindrical surface:* At all points, $\vec{E}$ is perpendicular to surface and has same magnitude. *End faces:* $\vec{E}$ is parallel to face, so flux is zero.
Planar (infinite charged sheet)	$\vec{E}$ is uniform and perpendicular to plane.	Cylinder or box perpendicular to plane 	*Surface perpendicular to plane:* $\vec{E}$ is parallel to face, so flux is zero. *Faces parallel to plane:* At all points, $\vec{E}$ is perpendicular to surface and has same magnitude.

QUANTITATIVE TOOLS

As you may imagine, this so-called *direct integration* is no simple matter (see Section 13.8 for a similar integral), even though the result, which we derived qualitatively in Section 24.4, is surprisingly simple. Taking advantage of the symmetry of the problem, however, we can use Gauss's law to calculate the answer in just two steps.

We begin by drawing a concentric spherical Gaussian surface of radius $r > R$ around the shell (**Figure 24.28*a***). According to Gauss's law, the flux through the Gaussian surface is equal to the enclosed charge divided by ϵ_0:

$$\Phi_E = \frac{q_{\text{enc}}}{\epsilon_0} = \frac{q}{\epsilon_0}. \tag{24.10}$$

In addition, we know that because of the spherical symmetry the electric field has the same magnitude E at each position on the Gaussian surface and the field is perpendicular to the surface. Because the electric field is perpendicular, we have $\vec{E} \cdot d\vec{A} = E\,dA$, and because E has the same value everywhere on the Gaussian surface, we can pull the electric field out of the integral in Eq. 24.5:

$$\Phi_E = \oint \vec{E} \cdot d\vec{A} = \oint E\,dA = E \oint dA = EA, \tag{24.11}$$

where A is the area of the spherical Gaussian surface:

$$A = 4\pi r^2. \tag{24.12}$$

Substituting Eq. 24.12 into Eq. 24.11, we obtain

$$\Phi_E = 4\pi r^2 E. \tag{24.13}$$

Combining Eqs. 24.10 and 24.13, we obtain

$$4\pi r^2 E = \frac{q}{\epsilon_0} \tag{24.14}$$

or

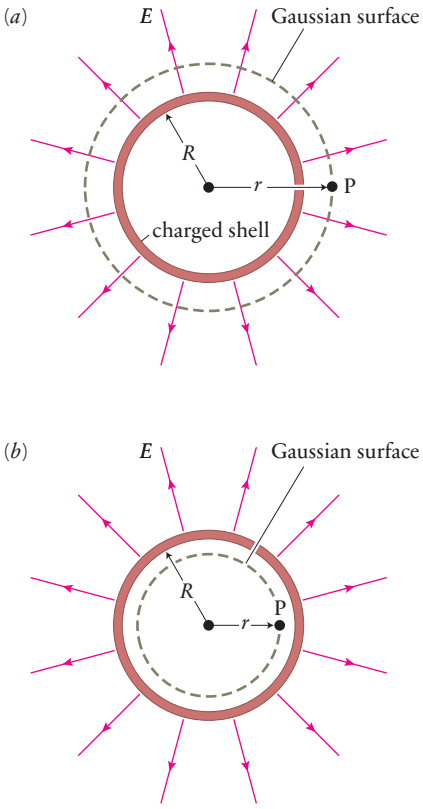
$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} = k\frac{q}{r^2}. \tag{24.15}$$

This is exactly the magnitude of the electric field due to a particle carrying a charge q located at the center of the shell, as we concluded in Example 23.7 for a solid sphere.

We can also use Gauss's law to determine the electric field inside the shell by drawing a concentric spherical Gaussian surface with radius $r < R$ (Figure 24.28*b*). For this surface the enclosed charge is zero, $q_{\text{enc}} = 0$, so the right side of Eq. 24.14 becomes zero. Consequently, the electric field inside the uniformly charged spherical shell must be zero.

Note that our calculation did not involve working out any integrals, even though Eq. 24.5 does contain an integral—the symmetry of the problem allows us to bypass the integration. The Procedure box on page 864 shows how to calculate the electric field using Gauss's law for charge distributions that exhibit one of the symmetries listed in Figure 24.27. In the next three exercises we apply this procedure to calculate the electric field of a number of different charge distributions.

Figure 24.28 Applying Gauss's law to a charged spherical shell.



PROCEDURE: Calculating the electric field using Gauss's law

Gauss's law allows you to calculate the electric field for charge distributions that exhibit spherical, cylindrical, or planar symmetry without having to carry out any integrations.

1. Identify the symmetry of the charge distribution. This symmetry determines the general pattern of the electric field and the type of Gaussian surface you should use (see Figure 24.27).
2. Sketch the charge distribution and the electric field by drawing a number of field lines, remembering that the field lines start on positively charged objects and end on negatively charged ones. A two-dimensional drawing should suffice.
3. Draw a Gaussian surface such that the electric field is either parallel or perpendicular (and constant) to each face of the surface. If the charge distribution divides space into distinct regions, draw a Gaussian surface in each region where you wish to calculate the electric field.
4. For each Gaussian surface determine the charge q_{enc} enclosed by the surface.
5. For each Gaussian surface calculate the electric flux Φ_E through the surface. Express the electric flux in terms of the unknown electric field E.
6. Use Gauss's law (Eq. 24.8) to relate q_{enc} and Φ_E and solve for E.

You can use the same general approach to determine the charge carried by a charge distribution given the electric field of a charge distribution exhibiting one of the three symmetries in Figure 24.27. Follow the same procedure, but in steps 4–6, express q_{enc} in terms of the unknown charge q and solve for q.

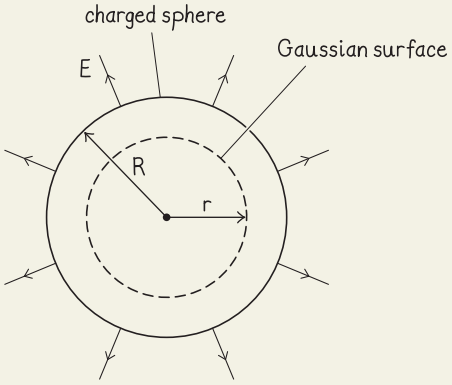
EXERCISE 24.6 Electric field inside uniformly charged sphere

Consider a charged sphere of radius R carrying a positive charge q that is uniformly distributed over the volume of the sphere. What is the magnitude of the electric field a radial distance $r < R$ from the center of the sphere?

SOLUTION The sphere has spherical symmetry, so I know that the field must point radially outward in all directions. I therefore draw a concentric spherical Gaussian surface with a radius $r < R$ (**Figure 24.29**). Because the sphere carries a uniformly distributed charge q, the amount of charge enclosed by the Gaussian surface is determined by the ratio of the volumes of the Gaussian surface and the charged sphere:

$$q_{\text{enc}} = \frac{\frac{4}{3}\pi r^3}{\frac{4}{3}\pi R^3} q = \frac{r^3}{R^3} q. \tag{1}$$

Figure 24.29



The electric flux is given by the product of the magnitude of the electric field $E(r)$ and the surface area A of the Gaussian surface (Eq. 24.13):

$$\Phi_E = 4\pi r^2 E. \tag{2}$$

Substituting Eqs. 1 and 2 into Eq. 24.8, I obtain

$$4\pi r^2 E = \frac{r^3}{R^3}\frac{q}{\epsilon_0}$$

or

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{R^3} r = k\frac{q}{R^3} r, \checkmark$$

the same result I obtained in Checkpoint 23.15.

24.22 What is the electric field outside a solid sphere carrying a charge $+q$ uniformly distributed throughout its volume?

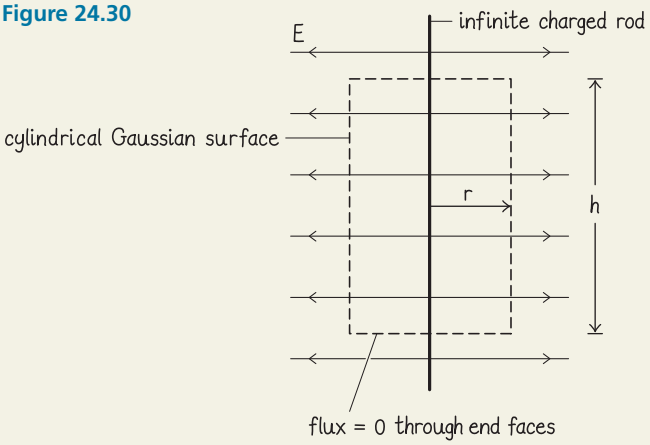
EXERCISE 24.7 Electric field of an infinitely long charged thin rod

What is the electric field magnitude a radial distance r from the central length axis of an infinitely long thin rod carrying a positive charge per unit length λ?

SOLUTION An infinitely long rod has cylindrical symmetry. I assume the rod's diameter is vanishingly small. From the symmetry, I know that the electric field points radially outward (see Section 24.4). I therefore make a sketch showing the rod and a few representative field lines. I then draw a cylindrical Gaussian surface of radius r and height h around the rod (**Figure 24.30**). The cylinder encloses a length h of the rod, so the enclosed charge is

$$q_{\text{enc}} = \lambda h. \tag{1}$$

Figure 24.30



The electric flux through the top and bottom faces of the Gaussian surface is zero because the electric field is parallel to those faces. I also know that symmetry requires the electric field to have the same magnitude E at each position on the cylindrical part of the surface. I can therefore pull the electric field out of the integral. The electric flux through that part of the surface is

$$\Phi_E = \int_{\text{cyl surface}} \vec{E} \cdot d\vec{A} = \int E\, dA = E \int dA. \tag{2}$$

The area of the cylindrical surface is equal to the circumference of the cylinder, $2\pi r$, times its height h, so Eq. 2 becomes

$$\Phi_E = E(2\pi r h). \tag{3}$$

Substituting Eqs. 1 and 3 into Eq. 24.8, I obtain

$$E(2\pi r h) = \frac{\lambda h}{\epsilon_0}$$

or

$$E = \frac{\lambda}{2\pi\epsilon_0 r} = \frac{2k\lambda}{r}. \checkmark$$

This is the same result I obtained by direct integration in Example 23.4 for a finite charged rod in the limit that I am close to the rod (which makes the rod appear infinitely long). The direct integration, however, took almost two pages of work!

24.23 The direct integration procedure also yields an expression for a rod of *finite* length (see Example 23.4). Can you use Gauss's law to derive this expression as well? Why or why not?

EXERCISE 24.8 Electric field of an infinite charged sheet

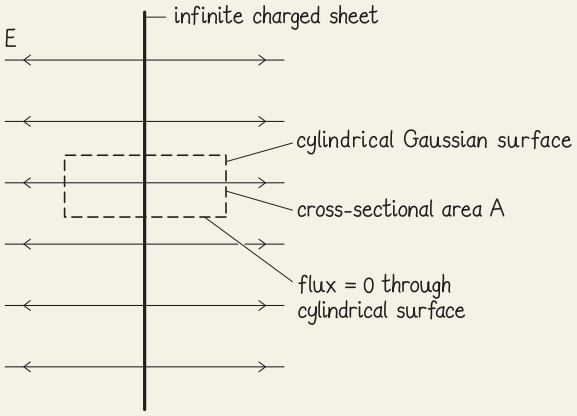
What is the electric field a distance d from a thin, infinite nonconducting sheet with a uniform positive surface charge density σ?

SOLUTION An infinite sheet has planar symmetry. From the symmetry I know that the electric field points away from the sheet and that the magnitude of the electric field is the same everywhere (see Section 24.4). I make a sketch of the sheet and the electric field and then draw a Gaussian surface in the form of a cylinder that straddles the sheet (**Figure 24.31**). If the cross section of the cylinder has area A, then the cylinder encloses a piece of the sheet of area A and the enclosed charge is

$$q_{\text{enc}} = \sigma A. \tag{1}$$

The electric flux through the cylindrical part of the Gaussian surface is zero because the field lines are parallel to that surface. The field is perpendicular to the two ends and

Figure 24.31



(Continued)

points outward, however, so the electric flux through those ends is the product of the area of each end, A, and the magnitude of the electric field:

$$\Phi_E = 2EA. \qquad (2)$$

Substituting Eqs. 1 and 2 into Eq. 24.8, I get

$$2EA = \frac{\sigma A}{\epsilon_0}$$

or

$$E = \frac{\sigma}{2\epsilon_0}. \checkmark$$

Because $1/(2\epsilon_0) = 2\pi k$, I can also write this as $E = 2\pi k\sigma$, which is the same result I obtained by direct integration in Example 23.6 for a uniformly charged disk in the limit that I am very close to the disk (which makes the disk appear like an infinite sheet).

The situation is a little different for a *conducting* plate. Consider, for example, the infinite charged conducting plate shown in **Figure 24.32a**. As we saw in Section 24.5, any charge resides on the outside surfaces of the conducting object, so we have to consider the charge on *both* surfaces. If the surface charge density on the plate is σ and we use the same cylindrical Gaussian surface as we did for the nonconducting sheet, then the enclosed charge is not σA but $2\sigma A$ (σA for each of the two surfaces). Then Gauss's law yields

$$\Phi_E = \frac{q_{\text{enc}}}{\epsilon_0} = \frac{2\sigma A}{\epsilon_0}. \qquad (24.16)$$

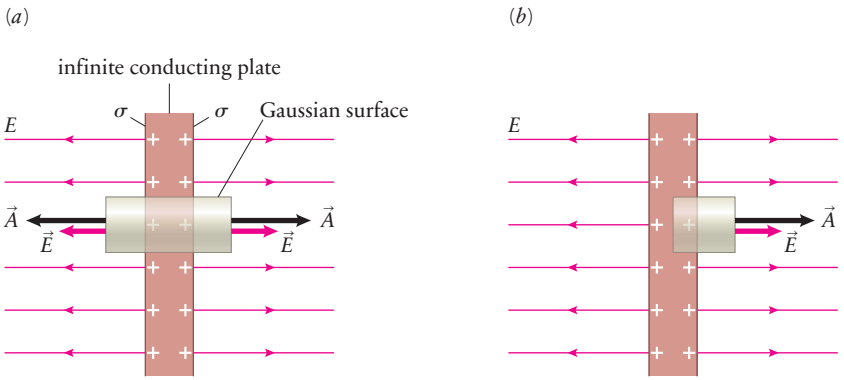
Substituting Eq. 24.16 into Eq. 2 of Exercise 24.8 yields

$$E = \frac{\sigma}{\epsilon_0} \quad \text{(infinite conducting plate)}. \qquad (24.17)$$

Alternatively, you can choose a cylindrical Gaussian surface that has one end embedded in the plate (Figure 24.32*b*). In that case only one of the two surfaces is enclosed and so the enclosed charge is σA. However, now the electric flux through the left end is zero because $E = 0$ inside the bulk of the plate, so the electric flux through the Gaussian surface is not $2EA$ but EA. Substituting these values into Gauss's law yields again Eq. 24.17.

24.24 (*a*) A very large metal plate of surface area A carries a positive charge q. What is the surface charge density of the plate? What is the magnitude of the field created by the plate? (*b*) A very large, thin nonconducting sheet of surface area A carries a fixed, uniformly distributed positive charge q. What is the surface charge density of the sheet? What is the magnitude of the field created by the sheet?

Figure 24.32 Applying Gauss's law to an infinite charged conducting plate.



QUANTITATIVE TOOLS

Chapter Summary

Electric field lines and electric flux (Sections 24.1–24.3, 24.6)

Concepts **Electric field lines** let us visualize the electric field around any charge distribution. The lines are drawn so that at any point the electric field $\vec{E}$ is tangent to the lines.

The number of electric field lines that emanate from a positively charged object or terminate on a negatively charged object is proportional to the charge carried by the object.

The **field line density** at a given position in an electric field is the number of field lines per unit area that cross a surface perpendicular to the field lines at that position. At every position in a field line diagram, the magnitude of the electric field is proportional to the field line density at that position.

The electric flux through a closed surface depends only on the amount of charge enclosed by the surface. The flux through a closed surface due to charge carriers outside the surface is always zero.

Quantitative Tools The *area vector* $\vec{A}$ associated with a flat surface area has a magnitude A equal to the surface area and is directed normal to the plane of the area.

The **electric flux** Φ_E ($\mathrm{N \cdot m^2/C}$) through a surface due to an electric field $\vec{E}$ is

$$\Phi_E = \begin{cases} \int \vec{E} \cdot d\vec{A} & (24.4) \\ \oint \vec{E} \cdot d\vec{A} \quad \text{(closed surface)}, & (24.5) \end{cases}$$

where $d\vec{A}$ is an infinitesimal surface area segment and $\vec{E}$ is the electric field at that location. If the surface is closed, the direction of $d\vec{A}$ is chosen to point outward.

Gauss's law (Sections 24.4, 24.7)

Concepts A **Gaussian surface** is any closed surface used to apply Gauss's law. When using Gauss's law to determine an electric field, we choose a Gaussian surface such that the electric field is the same (and possibly zero) everywhere along as many regions of the surface as possible.

The electric field outside a uniformly charged spherical shell is the same as the electric field due to a particle that carries an equal charge and is located at the shell center. The shell's contribution to the electric field is zero everywhere inside the shell.

Quantitative Tools **Gauss's law:** The electric flux through a Gaussian surface is

$$\Phi_E = \oint \vec{E} \cdot d\vec{A} = \frac{q_{\text{enc}}}{\epsilon_0}, \tag{24.8}$$

where q_{enc} (the **enclosed charge**) is the sum of all the charge inside the surface, and

$$\epsilon_0 \equiv \frac{1}{4\pi k} = 8.85 \times 10^{-12}\ \mathrm{C^2/(N \cdot m^2)} \tag{24.7}$$

is the **permittivity constant** ($k = 9.0 \times 10^9\ \mathrm{N \cdot m^2/C^2}$, Eq. 22.5).

Calculating electric fields (Sections 24.5, 24.8)

Concepts A Gaussian surface used to determine an electric field should have the same symmetry (spherical, cylindrical, planar) as the symmetry of the charge distribution producing the field. See Figure 24.27 and the Procedure box on page 864.

Quantitative Tools The magnitude of the electric field a distance r from the center of a sphere that has radius $R \geq r$ and contains charge of magnitude q distributed uniformly throughout the sphere volume is

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{R^3}r = k\frac{q}{R^3}r.$$

The magnitude of the electric field a distance r from an infinite rod having a uniform linear charge density λ is

$$E = \frac{\lambda}{2\pi\epsilon_0 r} = \frac{2k\lambda}{r}.$$

A system in which the distribution of charge does not change is in **electrostatic equilibrium.** The electric field inside a conducting object that is in electrostatic equilibrium is zero. Any surplus charge placed on the object arranges itself at the surface of the object, and the electric field at the surface is perpendicular to the surface. A metal reaches electrostatic equilibrium so quickly that this rearrangement is virtually instantaneous.

The magnitude of the electric field produced by a nonconducting sheet having a uniform surface charge density σ is

$$E = \frac{\sigma}{2\epsilon_0}.$$

The magnitude of the electric field produced by an infinite conducting plate having a uniform surface charge density σ is

$$E = \frac{\sigma}{\epsilon_0}. \tag{24.17}$$

Questions and Problems

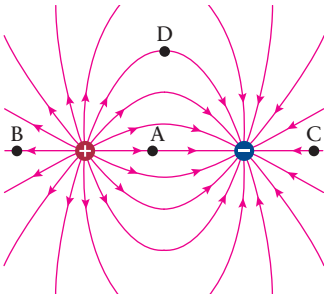
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

24.1 Electric field lines

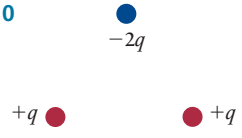
1. Suppose a charged particle is located at the origin. What is the direction of the electric field at the point (0.6, 1.2) if the particle is (*a*) positively charged and (*b*) negatively charged? •
2. During an airplane flight let it gets positively charges due to the friction with air and the dust particles. In that situation the flux of electric field into the aircraft is zero, positive or negative? Justify your answer. •
3. Two positive charges having magnitude +q each are separated by a distance 6 mm apart. Compute the electric field at a point located 3 mm: (*a*) on the line joining of the two charges and (*b*) not on the line joining of the two charges. •
4. You and a friend are asked to draw the two-dimensional electric field line pattern for two charged objects located near each other. The charge on the first object is $+2q$ (with q positive), and the charge on the second object is $-q$. You show 32 field lines emanating from the first charge and 16 field lines terminating on the second charge. Your friend shows 24 field lines emanating from the first object and 12 terminating on the second object. Which diagram is correct? •
5. A positively charged test particle is released from rest in a uniform electric field. The only force experienced by the particle is the electric force in that field. (*a*) Describe the trajectory the particle follows. (*b*) How does the trajectory change, if at all, if the particle is given an initial velocity in the direction of the electric field? (*c*) How does the trajectory change if the particle has an initial velocity that makes an angle $\theta \neq 0$ with the electric field? (*d*) Repeat parts *a*–*c* for a negatively charged test particle. ••
6. A certain field line diagram illustrates the electric field due to three particles that carry charges 6.0 μC, -2.0 μC, and -4.0 μC. If twelve field lines emanate from the positively charged particle, how many field lines terminate on each of the negatively charged particles? ••
7. Figure P24.7 shows the electric field produced by an electric dipole. How would a positively charged test particle begin to move when released from rest (*a*) at point A, (*b*) at point B, (*c*) at point C, and (*d*) at point D? Assume that no other objects exert forces on the test particle. (*e*) How does your answer to part *a* change if the test particle has an initial velocity that makes some angle $\theta \neq 0$ with the electric field at A? ••

Figure P24.7



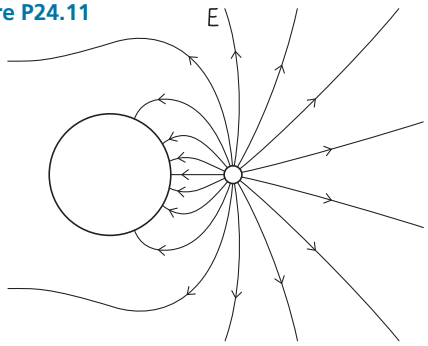
8. Two particles of charge $+q$ are located at opposite ends of one diagonal of a square, and two particles of charge $-q$ are located at opposite ends of the square's other diagonal. Sketch the pattern of electric field lines created, showing eight lines per particle in the plane that contains the square. ••
9. Assume that a test particle is released from rest in an electric field and experiences no forces other than the electric force exerted by that electric field. Does an electric field line passing through the particle's initial position represent the particle's trajectory over time? If not, illustrate with an example. ••
10. Draw several field lines surrounding the three charged particles shown in Figure P24.10. ••

Figure P24.10



11. Two classmates sitting next to you have made an electric field diagram. You are looking at it (not during a test) to see if they did it correctly, but part of the diagram is covered by a quarter, as shown in Figure P24.11. If your classmates made the diagram correctly, what can you say about the region covered by the quarter? State all the information you can determine about any charged particles in that region. ••

Figure P24.11

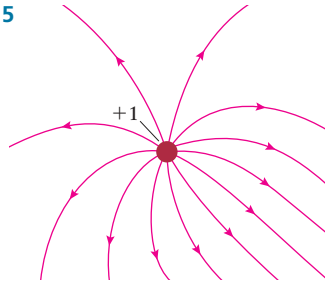


12. Outline an algorithm for a computer program that traces out the electric field lines surrounding a charged object. Assume that functions have already been provided to draw the charged object itself and to draw line segments between pairs of points that your program chooses. •••
13. There are many parallels between gravitational phenomena and electrical phenomena: One can think of a test object with mass in a gravitational field as analogous to a test object with positive charge in an electric field. (*a*) Draw the gravitational field line pattern for the Earth-Moon system, showing arrows on the lines to indicate direction. (*b*) Now consider a system made up of two charged objects. Which type of charge (positive or negative) must each object carry in order to have an electric field line pattern match the gravitational field line pattern that you drew in part *a*? (*c*) Do you see any differences between electrical and gravitational phenomena? •••

24.2 Field line density

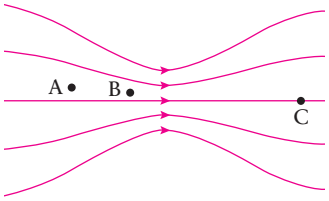
14. A small charged ball is suspended at the center of a spherical balloon that is nestled snugly inside a cubical cardboard box. On one side, the balloon touches the wall of the box. (*a*) At this point of contact, is the electric field magnitude on the balloon surface equal to the electric field magnitude on the box surface? Assume neither object is polarized. (*b*) At this point of contact, is the number of field lines per unit of balloon surface area equal to the number of field lines per unit of box surface area? •
15. Where in Figure P24.15 other than around the point charge is the electric field greatest in magnitude? •

Figure P24.15



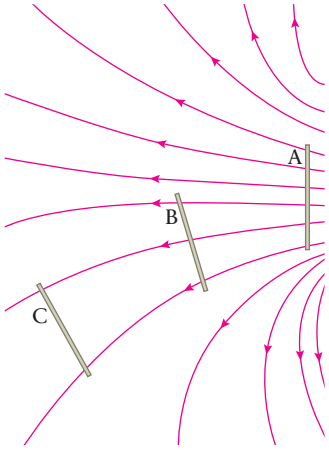
16. Consider 500 uniform electric field lines that pass through a hollow tube of length 10 m and cross-sectional area 1m^2. If the tube is cut at the middle in such way the new cross-section makes an angle of 60° with the electric field lines, then compare the area of the two cross-sections. •
17. Consider the pattern of five field lines shown in Figure P24.17. Rank positions A, B, and C according to the magnitude of their electric fields, from greatest to smallest. •

Figure P24.17



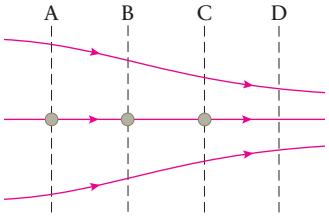
18. Make a geometric argument, based on how the field lines spread out into space, that the magnitude of the electric field surrounding a long, straight, charged wire is proportional to $1/r$, where r is the radial distance from the long axis of the wire. ••
19. An aluminum tube of length 8 m has been uniformly charged with 10 μC of charge. Considering symmetry, estimate the electric field at the geometric center of the tube lying on the axial position. ••
20. Figure P24.20 shows electric field lines due to a distribution of charge (not shown). The labeled regions of space are all equal in size. In which of the labeled regions, A, B, and C, is (*a*) the electric field greatest and (*b*) the electric field smallest? (*c*) How much greater is the largest labeled electric field than the smallest nonzero labeled electric field? ••

Figure P24.20



21. Identical charged pellets are released from rest from equidistant starting lines A, B, and C in Figure P24.21, in a region of nonuniform electric field. (*a*) Rank the pellets according to the time interval it takes the pellet on line A to cross line B, the pellet on B to cross C, and the pellet on C to cross D, smallest interval first. (*b*) How does your answer to part *a* change if each pellet is given an identical initial velocity directed upward along lines A, B, and C? ••

Figure P24.21

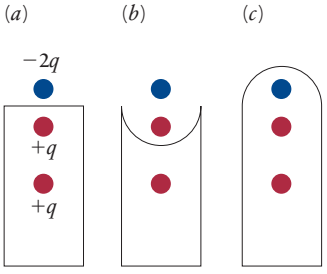


22. You decide to make an electric field line diagram in two dimensions, with 16 lines per coulomb of charge. (*a*) How many lines should you draw surrounding a particle that carries +1.5 C of charge? Should the lines enter or leave that particle? (*b*) Answer the same questions for a particle that carries −0.375 C of charge. (*c*) Could you represent a particle that carries +0.8 C in this diagram? ••
23. Two flat surfaces are near two different charged objects, and the surfaces are identical in size. Surface A has N field lines passing perpendicularly through it. Surface B has only $N/2$ field lines passing through it, and they pass through at an angle 72° from normal. Is the field line density greater near surface A or surface B, and by how much? •••

24.3 Closed surfaces

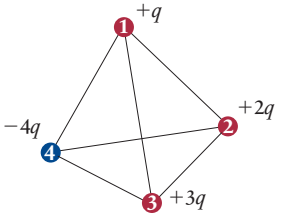
24. A negative charge is located inside the swim ring. Is there any location on the surface of the swim ring where the electric field line flux is positive? •
25. Closed surface A is a sphere of radius R, closed surface B is a sphere of radius $2R$, and closed surface C is a cube with side length R. At the geometric center of each closed surface is a small ball that is completely enclosed by the surface and carries electrical charge $+q$. Assuming there are no other charged objects inside any of the closed surfaces, rank the surfaces according to the electric field line flux through them, largest first. •
26. Three charged pellets are arranged along a line. Three hypothetical closed surfaces are drawn around this set of three pellets. The three surfaces have identical cylindrical walls but different top caps (Figure P24.26). Rank the closed surfaces from smallest to greatest in terms of the electric field line flux through them. •

Figure P24.26



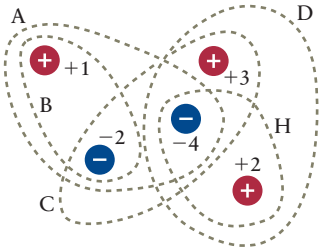
27. In Figure P24.27, assume that a Gaussian surface drawn to surround only sphere 1 would have an electric field line flux of +4. Sketch Gaussian surfaces that contain sphere 1 plus one or more additional spheres and have field line fluxes of (*a*) +24, (*b*) −4, and (*c*) +8. ••

Figure P24.27



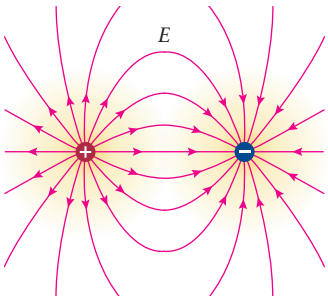
28. In the electric field line diagram for the charge arrangement shown in Figure P24.28, 12 field lines emanate from the object of charge +1 C. What is the field line flux through closed surfaces (*a*) A, (*b*) B, (*c*) C, (*d*) D, and (*e*) H? ••

Figure P24.28



29. In a certain field line diagram, a single electron is drawn that has four field lines terminating on it. If a closed surface in another part of this diagram is shown to have 16 field lines coming out of it, what can be said about the number of charge carriers inside the surface and the type of charge they carry? ••
30. For the dipole in Figure P24.30, can you draw a two-dimensional closed surface through which the field line flux is (*a*) zero, (*b*) +16, (*c*) −16, and (*d*) +3? ••

Figure P24.30



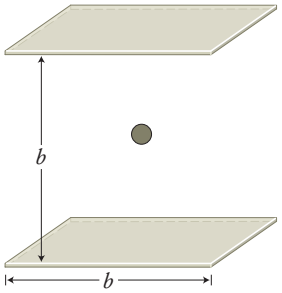
31. Make an argument based on Gauss's law that a charged object cannot be held in stable translational equilibrium by electric fields alone. For a start, assume there is no other charged object in the region of interest even though the external electric field can be due to charged objects outside this region. •••
32. The chapter explains that the charge inside a closed surface is proportional to the flux through that surface. To illustrate why the surface must be closed, draw an example of a charge distribution and a surface for which the flux through the surface would be considerably altered if the surface had any holes (unclosed regions of the surface through which flux is not counted). •••

24.4 Symmetry and Gaussian surfaces

33. (*a*) Can you draw a Gaussian surface that offers a simple way to compute the electric field surrounding a charged dipole? (*b*) If not, does this mean that the relationship between field line flux and enclosed charge and the relationship between field line density and magnitude of the electric field do not apply to charged dipoles? •
34. A charge of $+2q$ is enclosed inside a hollow copper sphere having a wall thickness of 5 mm. Compute the electric flux through the copper sphere. •

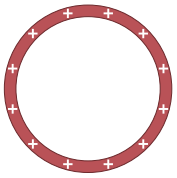
35. For a charged particle placed at the center of a spherical Gaussian surface, the field line flux through the surface is $+\Phi$. Suppose that the spherical Gaussian surface is replaced by a cubical Gaussian surface composed of six square plates. Square plates are positioned above and below the charged particle to form two sides of the cubical Gaussian surface that is centered on the charged particle, as shown in Figure P24.35. What is the field line flux through each plate? •

Figure P24.35



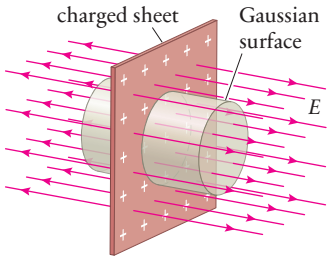
36. Figure P24.36 shows a cross section through an infinitely long, uniformly charged hollow cylinder. How does the magnitude of the electric field vary with the radial distance r from the cylinder's long central axis (a) inside the hole in the cylinder and (b) outside the cylinder? ••

Figure P24.36 Charged hollow cylinder



37. Figure P24.37 shows a cylindrical Gaussian surface straddling a charged sheet. Are there any other shapes that would be convenient for the Gaussian surface in this situation? ••

Figure P24.37

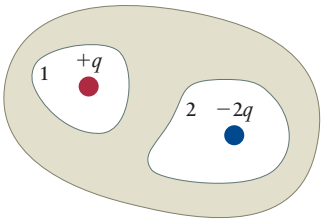


38. Suppose you want to determine the amount of charge on a spherical shell. You start out by surrounding the shell with a cylindrical Gaussian surface. Is it possible to use the flux everywhere on the surface of the cylinder to determine the charge on the spherical shell? Why isn't this recommended in the chapter? ••
39. The magnitude of the electric field a distance R away from a particle carrying charge q is E_0. Now consider this same quantity of charge q distributed uniformly throughout the volume of a nonconducting solid sphere of radius R. What is the electric field magnitude, in terms of E_0, (a) a distance $2R$ from the center of the sphere, (b) at the sphere surface, and (c) a distance $R/4$ from the sphere center? ••
40. You are given a uniformly charged spherical shell that has a radius R and carries a positive charge q. In order to measure the electric field magnitude inside, you drill a small hole that removes 0.01% of the shell's material and insert a probe into the shell. What are the direction and approximate magnitude of the electric field at the shell center? (Hint: The answer is small but not zero.) •••
41. You and your friend are working in a lab and wish to measure the surface charge density on a slab of metal. The slab is very long and very wide, but flat. You begin adding up the field line flux through a cylindrical Gaussian surface that passes through the slab as in Figure P24.37. Your friend begins adding up the field line flux through a cubical Gaussian surface that passes through the slab. Her cube is a little shorter than your cylinder, and it is wider than your cylinder. Will the two of you still obtain the same result for the surface charge density? If so, why? If not, what needs to be changed about your friend's cube? •••
42. You have a hollow spherical shell of radius R that carries a positive charge. The shell is very thin. You make many Gaussian surfaces concentric with the shell to determine the electric field line density at various distances from the center. Qualitatively plot the electric field line density as a function of the distance from the center of the shell. Show regions inside the shell as well as outside. •••

24.5 Charged conducting objects

43. The two-cavity metal object in Figure P24.43 is electrically neutral, but each cavity contains a charged particle as shown. What are (a) the charge on the surface of each cavity and (b) the charge on the outer surface of the object? •

Figure P24.43



44. Some charged particles are suspended inside a hollow metal object by supports made of a material that is an electrical insulator. The metal object has charge $+q$ distributed over its outer surface and charge $-2q$ distributed over its inner surface. (a) How much charge is suspended in the cavity? (b) What is the charge on the object? (Hint for part b: What is the electric field a very large distance r from the center of the object?) ••
45. Can a car protect the driver and the passengers from an external electric field? (b) Also, justify the need of keeping sophisticated electronic devices enclosed inside a metallic box. ••
46. A *coronal discharge* occurs when the electric field magnitude just above the surface of a conductor exceeds approximately 3×10^6 N/C, a magnitude strong enough to ionize atoms in the air by ripping electrons from them. Coronal discharge is most likely to occur near a sharp tip of a conductor. (This is why solder joints in

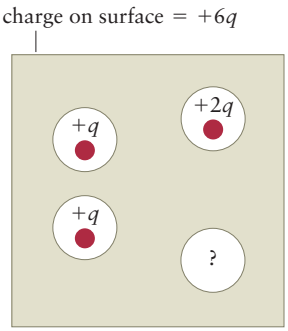
electronics must not have sharp protrusions and why it's easier to draw a spark from your fingertip than from your knee after scuffing your feet on a carpet.) (*a*) Make an argument based on the geometry of the field lines emerging from a conducting surface that charge is more likely to concentrate on a convex region of the surface than on a concave region. (*b*) Sketch a paring knife that has a surplus charge on it, and shade the regions where charge is most likely to concentrate on the surface. ••

47. Two concentric spherical metal shells are insulated from each other and from the surroundings. The inner shell carries a charge $+2q$, and the outer shell carries a charge $-q$. In electrostatic equilibrium, what is the charge on (*a*) the outer surface of the inner shell, (*b*) the inner surface of the outer shell, and (*c*) the outer surface of the outer shell? (*d*) The shells are then connected to each other by a metal wire. Do parts *a–c* for the situation obtained after electrostatic equilibrium is reached. ••

48. A small ball carrying a charge $-2q$ is placed at the center of a spherical metal shell that carries a charge $+q$. What are the sign and magnitude of the charge (*a*) on the inner surface of the shell and (*b*) on the outer surface of the shell? (*c*) Do your answers change if the ball is not at the center of the shell? If so, how? ••

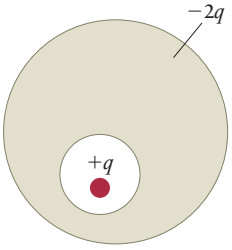
49. Figure P24.49 shows a neutral conducting block with four hollow chambers inside. A charge of $+6q$ has collected on the outer surface of the block. Inside three of the chambers are particles that carry charges $+q$, $-2q$, and $+3q$. Determine the charge inside the fourth chamber in terms of q. ••

Figure P24.49



50. A pellet carrying charge $+q$ is centered inside a spherical cavity that has been cut off-center in a solid metal sphere that carries charge $-2q$ (Figure P24.50). (*a*) How much charge resides on the inner surface of the cavity? Is this charge distributed uniformly or nonuniformly? (*b*) How much charge resides on the outer surface of the metal sphere? Is this charge distributed uniformly or nonuniformly? •••

Figure P24.50



51. A particle carries a charge of $2q$ and is located at the origin. The particle is surrounded by a conducting shell of inner radius R and thickness t. The shell carries a charge of $-3q$. Determine (*a*) the charge on the inner surface of the shell and (*b*) the charge on the outer surface of the shell. (*c*) Make a qualitative plot of the electric field line density as a function of distance from the origin. Be sure to show the field line density in the hollow region, inside the material of the shell, and outside. •••

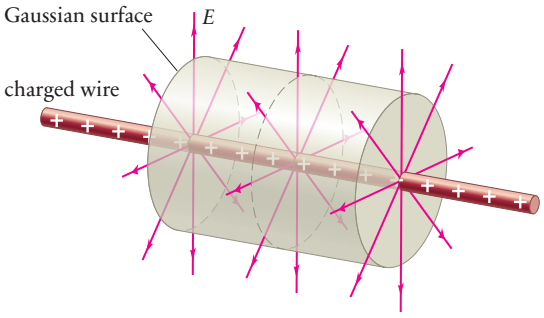
24.6 Electric flux

52. A charge of 35 nC is placed at the center of the rugby ball. Calculate the net flux through the balls surface. Comment on the location on the ball's surface where the electric field is minimum. •

53. A charge of 50 nC is located at the center of a hollow sphere of radius 10 cm. For what radius of the sphere will the electric field line density over the surface be: (*a*) doubled and (*b*) halved? •

54. Figure P24.54 shows a cylindrical Gaussian surface enclosing a segment of a long charged wire. Consider using a spherical Gaussian surface to enclose the same segment of the wire. (*a*) Which surface, cylindrical or spherical, has the greater electric flux through it? (*b*) Which surface is more appropriate for calculating the electric flux? Justify your answer. •

Figure P24.54

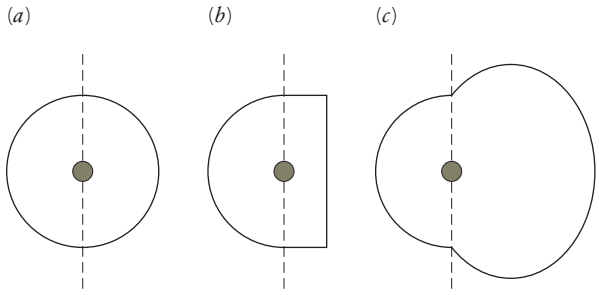


55. A particle carries a charge of 10 μC. Calculate the electric flux through single face if the charge is placed at the geometric center of (*a*) an octahedron and (*b*) a dodecahedron. •

56. Use a symmetry argument to determine the electric flux through a Gaussian surface that encloses both charged particles that form an electric dipole. Does the value calculated for the flux depend on the shape of the surface? ••

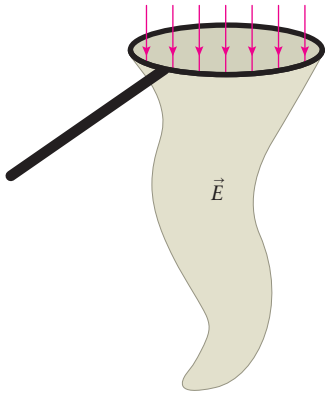
57. Figure P24.57 shows a charged particle surrounded by three different closed surfaces, (*a*), (*b*), and (*c*). In each case, the charge on the particle and the geometry of the left side (left of the dashed line) of the surface are identical. The closed surfaces have different geometries to the right of the dashed line. In which case is the flux through the right side of the surface the greatest? In which case is it the smallest? ••

Figure P24.57



58. A butterfly net hangs from a circular loop of diameter 400 mm. You hold the loop horizontally in a region where the electric field is 150 N/C downward, as shown in Figure P24.58. (*a*) What is the electric flux through the net? (*b*) Does the flux through the net increase or decrease if you rotate the loop so that it lies in a vertical plane? ●●

Figure P24.58



59. A particle that carries an unknown amount of charge is placed at the center of a hollow metal sphere of inner radius R_i and outer radius R_o. The magnitude of the electric field outside the sphere at any radial distance r from the center is $3kq/r^2$ (with $q > 0$) and is directed toward the center. The charge on the inner surface of the sphere is $-2q$. What are (*a*) the charge on the particle and (*b*) the charge on the sphere? ●●

60. A small conducting ball that carries a charge of 20.0 nC is located at the center of an electrically neutral hollow conducting spherical shell of inner radius 8 cm and outer radius 10 cm. What is the surface charge density on (*a*) the sphere's inner surface and (*b*) its outer surface? (*c*) Determine the electric flux through three concentric spherical Gaussian surfaces having radii 5 cm, 9 cm, and 11 cm, respectively. ●●

61. A uniformly charged sphere 5 cm in radius carries a charge of 8.0 nC. A thick hollow conducting spherical shell of inner radius 10 cm and outer radius 12 cm having a charge 4.0 nC is concentric with the sphere. (*a*) Calculate the surface charge density on the inner and outer radius of the shell. (*b*) Draw three concentric spherical Gaussian surfaces having radius 8 cm, 11 cm, and 14 cm. What is the electric flux through each of these surfaces? (*c*) Compute the electric flux through a spherical Gaussian surface concentric with the sphere, and having a radius of 2 cm. ●●●

24.7 Deriving Gauss's law

62. Three particles each carrying charge $+q$ are located at the corners of an equilateral triangle of side length a, and the triangle is centered on the origin of a Cartesian coordinate system. What is the electric flux through a spherical Gaussian surface centered at the origin if the radius of the Gaussian surface is (*a*) $a/4$ and (*b*) $2a$? ●

63. A particular closed surface has four sides. The electric flux is $+800\ \text{Nm}^2/\text{C}$ through side 1, $+300\ \text{Nm}^2/\text{C}$ through side 2, and $-1200\ \text{Nm}^2/\text{C}$ through side 3. (*a*) If no surplus charge is enclosed by the surface, what is the flux through side 4? (*b*) If the flux through side 4 is $-400\ \text{Nm}^2/\text{C}$, how much surplus charge is enclosed by the surface? ●

64. Consider a dipole located at the center of a cylindrical Gaussian surface such that the axis of the dipole aligns with the axis of the cylinder. (*a*) Find the total flux through the Gaussian surface. (*b*) Comment on the total flux through the Gaussian surface if the if the dipole is translated in such a way that the negative charge is moved out of the cylindrical surface. ●

65. Consider a rod of length ℓ that carries a uniform charge distribution. Is Gauss's law practical for calculating the electric field at a radial distance r from the midpoint of the rod's length (*a*) when $r \ll \ell$, (*b*) when $r = \ell$, and (*c*) when $r \gg \ell$? For each case in which your answer is yes, describe the shape of the most useful Gaussian surface. ●●

66. Any collection of N charged particles can be divided into N collections each containing one of the particles. Use the superposition property of electric fields to show that the electric flux through any Gaussian surface due to a collection of N charged particles is the sum of the fluxes through the same Gaussian surface produced by the N individual particles. ●●

67. A horizontal rectangular sheet of length ℓ and width w is positioned with its center a vertical distance $w/2$ below an infinitely long positively charged rod that has a uniform linear charge density λ. (*a*) If the sides of length ℓ are parallel to the rod's axis, what is the magnitude of the electric flux through the sheet? (*b*) If the vertical distance between the sheet center and the rod is reduced to $w/4$, does the flux magnitude increase, decrease, or stay the same? ●●

68. A cubical shell with edges of length a is positioned so that two adjacent sides of one face are coincident with the $+x$ and $+y$ axes of a rectangular coordinate system and the corner formed by these two sides is at the origin. (*a*) Sketch the electric field vectors in the region of the shell if an electric field of magnitude $E = bx^2$ directed along the $+x$ axis exists in this region. (*b*) How much charge is contained in the volume of the shell? ●●

69. A cubical Gaussian surface 30 mm on each side is centered on a particle that carries a charge of $+3.0\ \mu\text{C}$. (*a*) If you draw a field line diagram with four field lines per microcoulomb, how many lines pass through the Gaussian surface? In which direction are these lines? (*b*) Ignoring corners and edges, how many field lines pass through each face of the Gaussian surface? (*c*) What is the electric flux through the Gaussian surface? (*d*) What is the electric flux through one face? (*e*) Which of your answers change if the particle is moved 10 mm away from the center of the Gaussian surface? (You don't need to obtain new values for the answers that change, if any.) ●●

70. There is a Gauss's law for gravity analogous to Gauss's law for electricity. (*a*) If the electric flux through a closed surface is proportional to the enclosed charge, what is the gravitational flux proportional to? (*b*) Write an equation for Gauss's law for gravity. ●●

71. A dipole that carries charges $\pm q$ is aligned along the y axis of an xyz rectangular coordinate system, with the midpoint of the dipole at the origin. (*a*) Use Gauss's law to calculate the electric flux through the xz plane, which bisects the dipole and is perpendicular to the dipole moment. (Hint: Calculate the electric flux for each charge separately and apply the principle of superposition.) (*b*) Calculate the flux through the xz plane by integrating the electric field in Eq. 23.8 over the plane. (You'll need to generalize Eq. 23.8 to any point on the xz plane by making the substitution $x^2 \rightarrow r^2 = (x^2 + z^2)$.) (*c*) Which of the two techniques lets you obtain the flux through the xz plane more easily? ●●●

24.8 Applying Gauss's law

72. A positively charged hollow conducting sphere of radius 20 cm has a charge of 10 nC. If the same charge is distributed over the volume of a sphere having the radius 20 cm, compare the magnitude of the electric field at a distance of 1 m for both the hollow and solid spheres. Comment on the result. ●

73. A uniformly charged solid sphere of radius 10 cm has a charge of 5 nC. Determine (*a*) the volume charge density on the sphere, and the electric field, (*b*) 2 cm, and (*c*) 11 cm from the center of the sphere. ●

74. A positively charged thin cylindrical shell of length 50 cm and radius 1 mm has no end caps and a uniform surface charge density of 3×10^{-6} C/m^2. (*a*) What is the charge on the shell? Determine the electric field magnitude far from either end of the shell at (*b*) $r = 0.9$ mm and (*c*) $r = 1.1$ mm, where r is the radial distance from the long central axis of the shell. ●

75. An infinitely long positively charged cylindrical shell of radius a has a uniform linear charge density λ_a. An infinitely long cylindrical shell of radius $b > a$ is concentric with the first and has a uniform linear charge density λ_b. (*a*) Determine $\vec{E}$ in all regions. (*b*) If $\lambda_a = +5.0$ nC/m, which type of charge must the shell of radius b carry and what value must λ_b have in order for the electric field to be zero for radial distances $r > b$? (*c*) Sketch the electric field lines for this situation. ●●

76. Two parallel vertical sheets of infinite extent carry surplus charge and are 50 mm apart. Determine the magnitude and direction of the electric field in all regions of space (*a*) when each sheet has a uniform surface charge density $\sigma = +3.0$ nC/m^2 and (*b*) when $\sigma = +3.0$ nC/m^2 for the left sheet and $\sigma = -3.0$ nC/m^2 for the right sheet. (*c*) Draw a diagram showing the sheets and the electric field lines for the situations described in parts *a* and *b*. ●●

77. A long, thin, positively charged wire runs along the long central axis of a hollow conducting cylinder of length 10 m, inner radius 50 mm, and outer radius 70 mm. The wire has a uniform linear charge density of $+1.5$ μC/m, and the cylinder has zero net charge. (*a*) Ignoring the regions near the two ends, calculate the surface charge density on the cylinder's inner and outer surfaces. Again ignoring the ends, determine the electric field magnitude at (*b*) $r = 49$ mm, (*c*) $r = 51$ mm, and (*d*) $r = 100$ mm, where r is the radial distance from the wire. ●●

78. A positively charged nonconducting hollow sphere of outer radius R and inner radius $R/2$ has a uniform volume charge density ρ. Determine the magnitude and direction of the electric field as a function of r, the radial distance from the sphere center. ●●

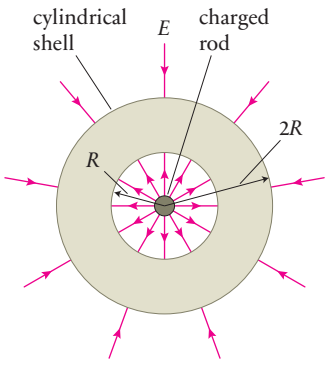
79. A positively charged nonconducting sphere of radius a has a uniform volume charge density ρ_0. It is snugly surrounded by a positively charged thick, nonconducting spherical shell of inner radius a and outer radius b. This thick shell has a volume charge density $\rho_0 r/a$ for $a < r \leq b$. Determine the electric field in all regions as a function of r, the radial distance from the center of the sphere. ●●

80. An infinitely long nonconducting solid cylinder of radius R has a nonuniform but cylindrically symmetrical charge distribution. The volume charge density is given by $\rho(r) = c/r$, where c is a positive constant having units C/m^2 and r is the radial distance from the long central axis of the cylinder. (*a*) What is the charge in a section of the cylinder of length ℓ? Write an expression for the electric field magnitude for (*b*) $r < R$ and (*c*) $r > R$. ●●

81. An infinitely long positively charged wire with a uniform linear charge density $+\lambda$ is parallel to the y axis of a Cartesian coordinate system and passes through the x axis at $x = -d$. At the origin, the electric field due to this charged wire has magnitude E_0 and is directed to the right along the positive x axis. A second infinite charged wire, parallel to the first and with a uniform linear charge density, passes through the x axis at $x = +3d$. The vector sum of the electric field at the origin due to the two lines of charge has magnitude $2E_0$. What is the linear charge density of the second charged wire? Obtain all the possible answers. ●●

82. Figure P24.82 shows a cross section through a system that consists of an infinitely long charged rod centered in a thick cylindrical shell of inner radius R and outer radius $2R$. The shell is made of a conducting material; the rod is made of a nonconducting material and has a uniform linear charge density. The electric field associated with the system is indicated by the field lines in the diagram. For each finite length ℓ of the system, there is a charge of $-q$ on the shell inner surface. For that same finite length ℓ, determine the charge on (*a*) the rod and (*b*) the shell outer surface. (*c*) In length ℓ, what is the charge on the shell? (*d*) What is the ratio of the surface charge density on the shell's inner surface to the surface charge density on the outer surface? ●●

Figure P24.82

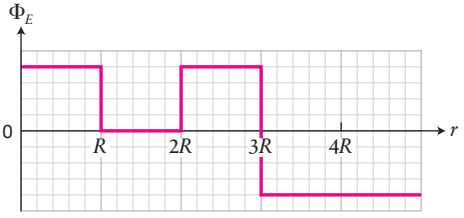


83. A positively charged solid non-conducting cylinder of length $\ell = 1$ m and radius $R = 2$ cm has a uniform volume charge density of 9×10^{-6} C/m^3. (*a*) What is the charge inside the cylinder? Avoiding the regions near the ends, determine the electric field magnitude at (*b*) $r = 1$ cm and (*c*) $r = 3$ cm, where r is the radial distance from the long central axis of the cylinder. ●●

84. Three nonconducting infinite sheets are parallel to each other. Each sheet has a uniform surface charge density. Sheet 1 is negatively charged with surface charge density $-\sigma$, and is defined to be the plane $x = 0$. Sheet 2 has an unknown surface charge density and is at a distance of 2.0 m to the right of sheet 1. Sheet 3, negatively charged with surface charge density -3σ, is located further to the right of Sheet 2 at a distance of 2.0 m from sheet 2. The electric field is zero at a point between Sheet 1 and Sheet 2, at a distance of 0.5 m from Sheet 1. (*a*) What is the surface charge density on sheet 2? What is the electric field at (*b*) $x = -1.0$ m, (*c*) $x = 3.0$ m, and (*d*) $x = 7.0$ m? ●●

85. A particle that carries positive charge $+q$ is located at the center of one or more concentric, thick or thin spherical shells. The shells are either made of a conducting material but are electrically neutral or made of a nonconducting material and have uniform surface charge densities. Figure P24.85 is the graph of the electric flux passing through spherical Gaussian surfaces centered on the particle, as a function of radial distance r from the particle. (*a*) What is the minimum number of spherical shells required to produce the graph? (*b*) Describe each shell completely, specifying its charge, its inner and outer radii, and whether it is made of a conducting or nonconducting material. ●●

Figure P24.85



86. Can the graph in Figure P24.85 be produced by a system that consists of only charged particles? If not, explain why not. If so, describe the system that contains the minimum number of charged particles. ●●

87. At what minimal radius along the perpendicular bisector of a wire of length 0.25 m carrying 30 nC of charge does the assumption of cylindrical symmetry in applying Gauss's law give an answer whose error exceeds 5%? ●●

88. An infinitely long positively charged wire carries a uniform linear charge density λ. A charged particle, carrying a charge q, is placed at a distance $2d$ from the wire. (*a*) What is the relation between q and λ so that the magnitude of the electric field at the midpoint between the particle and the wire is pointing toward the charge and has magnitude $2E$, where E is the magnitude of the electric field produced by the wire? (*b*) At what positions is the vector sum of the electric fields generated by the wire and by the particle equal to zero? (*c*) What is the magnitude of the electric field produced by the wire at the locations identified in part (*b*)? (Express your results in terms of E.) ●●

89. The origin of an x axis is placed at the center of a nonconducting solid sphere of radius R that carries a charge q_{sphere} distributed uniformly throughout its volume. A particle that carries an unknown charge q_{part} is located on the x axis at $x = +2R$. The magnitude of the electric field due to the sphere-particle combination is zero at $x = +R/4$ on the x axis. (*a*) What is q_{part} in terms of q_{sphere}? (*b*) At what other locations on the x axis is the electric field zero? ●●●

90. A particle that carries charge $+4q$ is located at the origin of an x axis, and a uniformly charged nonconducting solid sphere of radius R and carrying charge $+q$ is centered at $x = +6R$ on the x axis. (*a*) At what locations on the x axis is the electric field zero? (*b*) If the nonconducting sphere is replaced by a conducting sphere of the same radius and carrying the same charge, would the answers to part *a* change? ●●●

91. The nonuniform volume charge density inside a positively charged solid nonconducting sphere of radius R is $\rho(r) = \rho_0 r/R$, where r is the radial distance from the sphere center. (*a*) Calculate the electric field as a function of r for $r < R$ and $r > R$. Repeat for volume charge densities of (*b*) $\rho(r) = \rho_0(1 - \frac{1}{2}R/r)$ and (*c*) $\rho(r) = \rho_0(1 - R/r)$. ●●●

92. A positively charged nonconducting solid sphere of radius R has a nonuniform volume charge density given by ρ_0 for $r \le R/2$ and given by $2\rho_0(1 - r/R)$ for $R/2 \le r \le R$, where r is the radial distance from the sphere center. (*a*) Determine the charge q on the sphere in terms of ρ_0 and R. (*b*) Determine $\vec{E}$ in terms of ρ_0, r, and R. (*c*) Show that the electric field is continuous at the boundaries. ●●●

Additional Problems

93. In electrostatics, the electric field lines are always associated with source charges. If there is no electric field at a particular location in space, can we say that there are no charges surrounding that location? ●

94. Two charges $+q$, each having same mass m, were placed at rest at a distance of 10 cm and 20 cm above an infinite sheet that has been negatively charged with surface density σ. Compare the acceleration of the two charges. ●

95. A particle carrying charge q and having mass m is placed at rest at a distance d above an infinite sheet that has been charged to surface charge density σ. The sheet carries the same type of charge as the particle. What is the velocity of the particle when it is a distance $2d$ above the sheet? ●

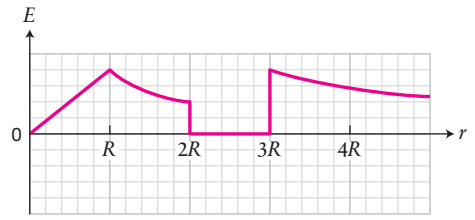
96. Infinite positively charged sheet 1 has uniform surface charge density $\sigma_1 = +4.0$ nC/m^2 and is located in the yz plane of a Cartesian coordinate system. Infinite negatively charged sheet 2 is parallel to sheet 1 and has uniform surface charge density $\sigma_2 = -8.0$ nC/m^2. Sheet 2 is 4 m to the right of sheet 1 along the $+x$ axis. (*a*) Determine the magnitude and direction of the electric field in all regions of space. (*b*) Draw a diagram showing the

locations of the two sheets, labeling the regions and drawing the field lines for each region. ●●

97. A conducting, thick spherical shell of inner radius 100 mm and outer radius 120 mm has no net charge. A particle carrying a 3.0-nC charge is placed at its center. (*a*) Draw the electric field line diagram for the situation, marking the location of all charge on your diagram. (*b*) Calculate the surface charge density at the inner and outer shell surfaces. (*c*) Determine the electric field as a function of *r*, the radial distance from the center of the shell. (*d*) Draw an electric field line diagram for the situation when the particle is displaced by 30 mm from the shell's center. ●●

98. A system is composed of two infinitely long concentric cylinders, an outer cylinder O and an inner cylinder I. Each is either conducting or nonconducting, and both have a uniform surface or volume charge density. Cylinder O is a shell, and cylinder I is either solid or a shell. The graph in Figure P24.98 shows the magnitude of the electric field as a function of the radial distance *r* from the long central axis of the concentric cylinders. At all values of *r*, the electric field either is zero or has a nonzero magnitude and is directed away from the long central axis. (*a*) Deduce whether the cylinders are conducting or nonconducting, and give the relevant radii *R* (inner and outer radii for O; either one radius or two for I depending on whether it is solid or a shell) in terms of the general radius *R* shown in the graph. (*b*) What is the ratio of the charge on cylinder I to the charge on cylinder O? ●●

Figure P24.98

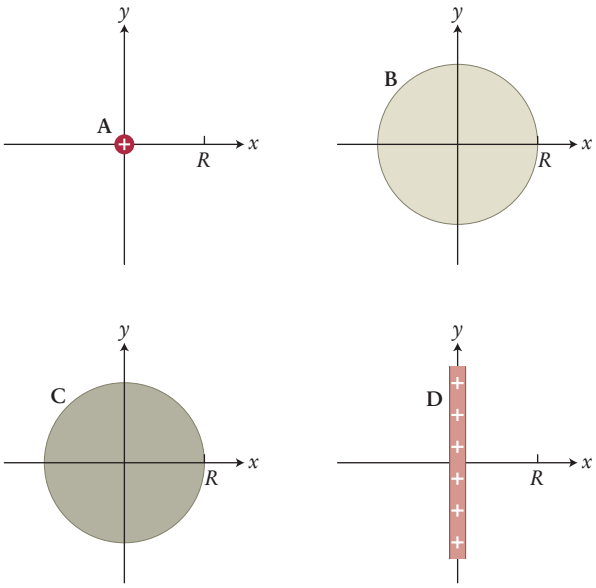


99. Near Earth's surface, the electric field has a magnitude of 150 N/C and points downward. Treating Earth's surface as a conducting sphere, calculate the surface charge density necessary to produce this electric field. ●●

100. An electron is fired toward a nonconducting infinite sheet carrying a uniformly distributed charge from 5.00 m away. The electron has an initial speed of 3×10^6 m/s along a line perpendicular to the sheet. The electron reverses its motion when it is 2.00 m away from the sheet. (*a*) What is the surface charge density on the sheet? (*b*) Given the same initial velocity, from what distance should the electron be fired if it is to just reach the sheet? ●●

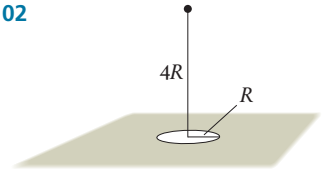
101. Figure P24.101 shows four charge distributions in an *xyz* coordinate system. A is a charged particle at the origin, B is a charged conducting solid sphere of radius *R* centered at the origin, C is a uniformly charged nonconducting solid sphere of radius *R* centered at the origin, and D is a nonconducting infinite sheet of uniformly distributed charge in the *yz* plane, shown in cross section. In each case the electric field at $x = R$ is the same. Rank the magnitudes of the electric fields for the four cases, largest first, at (a) $x = 2R$ and (*b*) $x = R/2$. ●●

Figure P24.101



102. A positively charged nonconducting infinite sheet that has a uniform surface charge density σ lies in an *xy* plane. As shown in Figure P24.102, a circular region of radius *R* has been cut out of the sheet. (*a*) What is the electric field magnitude at the center of the circle? (*b*) Use superposition to calculate the electric field, both magnitude and direction, at a distance $z = 4R$ above the center of the circle. ●●

Figure P24.102



103. A nonconducting positively charged slab 1 of thickness 2*a* has a nonuniform but symmetrical volume charge density $\rho_0 x/a$. The sheet is centered, at $x = 0$, on the *yz* plane of a Cartesian coordinate system. (*a*) Determine the magnitude and direction of the electric field in all regions of space as a function of *x*, the distance from the *yz* plane. (*b*) A nonconducting charged slab 2, of thickness 2*t*, is placed parallel to sheet 1 such that sheet 2 intersects the *x* axis at $x = d$. Sheet 2 is positively charged and has a uniform volume charge density ρ_0. Determine the electric field in all regions of space as a function of *x*. ●●●

104. A quantity of positive charge is distributed throughout the volume of a long nonconducting rod, which lies on the *z* axis and is centered at $z = 0$. The rod's radius *R* is much smaller than its length ℓ. For positions $r \ll \ell$ and $|z| \ll \ell$, calculate the electric fields inside and outside

the rod as a function of the radial distance r from the rod's long central axis if the volume charge density is (*a*) ρ_0 and (*b*) $\rho_0(1 - 2r/R)$. ●●●

105. You are a crewmember on a space station orbiting Earth. To investigate some physics you remember from college, you drill a small hole all the way through a large solid nonconducting sphere of radius R that has a negative charge q_s evenly distributed throughout its volume, making sure the hole goes through the sphere's center. You then go outside the station with the sphere and release a small pellet carrying positive charge $+q_p$ at one entrance of the hole. You are gratified to observe that the pellet oscillates with the frequency you calculated beforehand based on Gauss's law and your knowledge of simple harmonic motion. What is that frequency? ●●● CR

106. You need to create a uniform electric field in a small region of space, and all you have to work with is one large styrofoam ball that carries a uniform volume charge distribution ρ. You cut the ball into two sections and scoop a hemispherical hole out of each section so that putting the sections back together results in a spherical cavity *off-center* inside the ball. You know that the cavity can be described as a region composed of equal amounts of uniformly distributed positive and negative charge carriers, giving it a charge of zero. You suspect the principle of superposition might be useful. What is the electric field magnitude in the cavity? ●●● CR

107. Electrical power lines are often not insulated, and the magnitude of the electric field of the surrounding air is about 3×10^6 N/C. At greater electric field magnitudes, the air dissociates into charged particles, causing a breakdown and a big spark. You've been given a long sample of wire to test in the lab under electrostatic conditions, and you measure its diameter to be 16.6 mm. You have lots of equipment to produce a uniform linear charge density on the wire sample, and you figure you'd better build in a safety factor of 3 to account for humid days and so forth. What maximum charge density is safe? ●●● CR

108. An infinitely large positively charged nonconducting sheet 1 has uniform surface charge density $\sigma_1 = +130.0\ \text{nC/m}^2$ and is located in the xz plane of a Cartesian coordinate system. An infinitely large positively charged nonconducting sheet 2 has uniform surface charge density $\sigma_2 = +90.0\ \text{nC/m}^2$ and intersects the xz plane at the z axis, making an angle of 30° with sheet 1. (*a*) Draw a diagram showing the end view of the sheets in the xy plane. (*b*) Determine the expression for the electric field in the region between the sheets for positive values of x and y. Calculate the value of the electric field at (3 m, 1 m, 0). ●●●



25

Work and Energy in Electrostatics

CONCEPTS

25.1 Electric potential energy
25.2 Electrostatic work
25.3 Equipotentials

QUANTITATIVE TOOLS

25.4 Calculating work and energy in electrostatics
25.5 Potential difference
25.6 Electrostatic potentials of continuous charge distributions
25.7 Obtaining the electric field from the potential

As we saw in our study of mechanics, it is often easier to solve problems by using the concepts of energy and work than by using forces. In this chapter we study how to apply energy considerations to electric interactions. Because there are two types of charge—positive and negative—the energy changes associated with changes in charge configurations are a bit more complicated than those associated with changes in gravitational configurations. We first analyze the potential energy associated with a stationary charge distribution and then introduce a new quantity, *potential difference*, that is related to potential energy and that plays an important role in electronics because, unlike potential energy, it can be measured readily.

25.1 Electric potential energy

Figure 25.1 shows the energy changes that occur in closed systems of two charged objects. In Figure 25.1*a*, a positively charged particle is released from rest in the constant electric field of a massive stationary object carrying a negative charge. The attractive electric interaction between the particle and the object accelerates the particle toward the object (assuming that the only interaction is electric), and so the kinetic energy of the system increases. This increase in kinetic energy must be due to a decrease in **electric potential energy,** the potential energy associated with the relative positions of charged objects. As we can see, the electric potential energy of two oppositely charged objects decreases with decreasing separation between the two.

Figure 25.1*b* shows what happens in a system of two objects carrying like charges. In this case the particle is accelerated away from the object, and so the kinetic energy increases and the electric potential energy decreases with *increasing* separation between the two.

The situation depicted in Figure 25.1*a* is the electric equivalent of free fall. While all objects in free fall near Earth's surface experience the same acceleration, objects in electric fields experience *different* accelerations. Consider, for example, two particles with different masses in free fall near Earth's surface (**Figure 25.2*a*** on the next page). Even though the particle with the greater mass is subject to a greater gravitational force, its acceleration is the same as that of the particle with the smaller mass. The reason is that the particle's inertia (its resistance to acceleration) is equal to its mass (see Chapter 13). Now consider the situation illustrated in Figure 25.2*b*. Two particles carrying the same charge $+q$, but of different mass, are released near the surface of a massive negatively charged object. The electric forces exerted by the negatively charged object on the two particles are equal in magnitude, but because the masses of the two particles are different, the particles' accelerations are different as well.

25.1 Suppose both particles in Figure 25.2*b* are released from rest. Let $m_2 > m_1$ and consider only electric interactions. (*a*) How do their kinetic energies compare after they have both undergone the same displacement? (*b*) How do their momenta compare? (*c*) How do their kinetic energies and momenta compare at some fixed instant after they have been released? (*d*) How would you need to adjust the charges on the particles in order for the two particles to have the same acceleration upon release?

Changes in electric potential energy can also be associated with changes in the orientation of charged objects. Consider, for example, the situation illustrated

CONCEPTS

Figure 25.1 Energy diagrams for closed systems in which a positively charged particle is released from rest near a massive stationary object that carries a (*a*) negative or (*b*) positive charge.

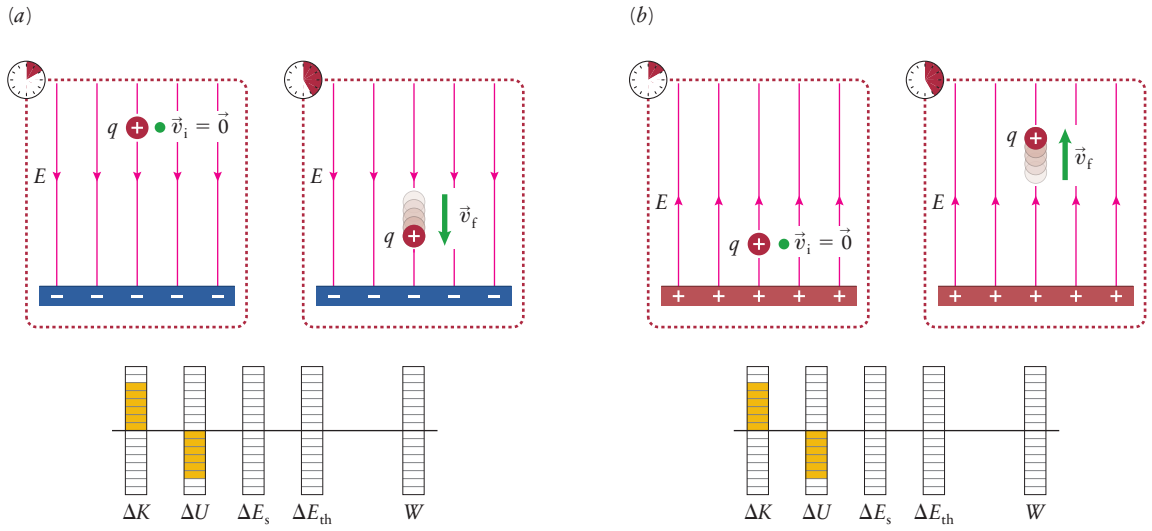
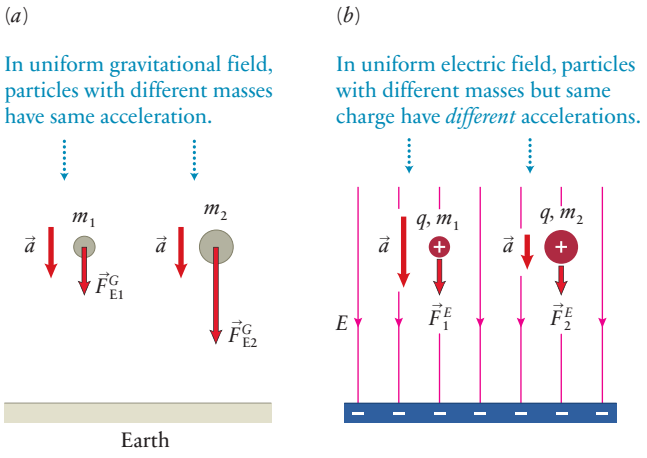


Figure 25.2 The free motion of charged particles in an electric field is different from the free fall of objects in a gravitational field.



in **Figure 25.3**. An electric dipole is held near the surface of a massive positively charged object. If the electric field of the massive object is uniform, then the dipole begins rotating as shown (see Sections 23.4 and 23.8). As it begins rotating, the dipole gains rotational kinetic energy. This means that the electric potential energy of the system must be changing as the orientation of the dipole changes. This change occurs because the positive side of the dipole gets farther away from the positively charged object, while the negative side gets closer to it; as we've seen, both of these motions correspond to a decrease of electric potential energy.

25.2 As the dipole in Figure 25.3 continues to rotate, it reaches the point where its axis is aligned with the electric field of the massive object. (*a*) What happens to the electric potential energy as the dipole moves beyond that point? (*b*) Describe the motion of the dipole beyond that point. (*c*) How would the motion of the dipole change if it were released with a different orientation from the one shown in Figure 25.3?

Figure 25.3 Energy diagram for a system in which a dipole is released from rest near a positively charged stationary object.

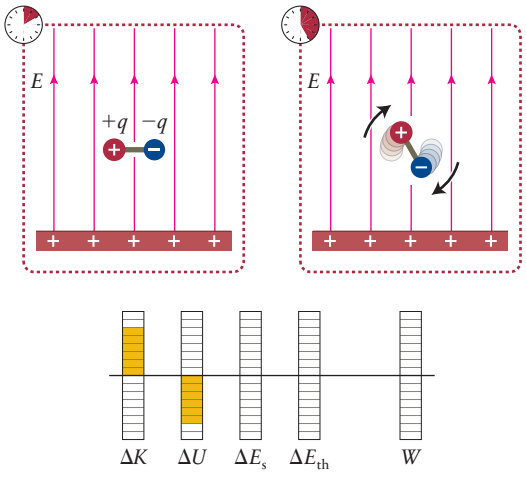
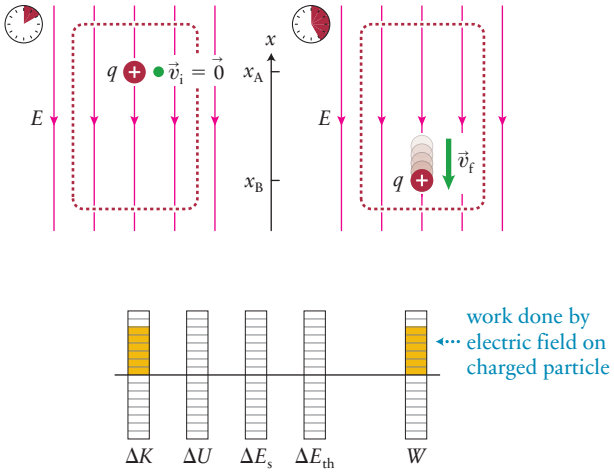


Figure 25.4 Energy diagram for a positively charged particle in the uniform electric field of a stationary, negatively charged object that is not part of the system.

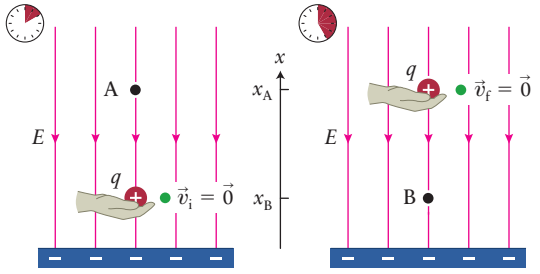


25.2 Electrostatic work

In general, we shall be considering the motion of a charged object through the constant electric field created by other stationary charged objects (such a field is sometimes called an *electrostatic field*). Therefore, our system—the charged object whose motion we are considering—is not closed. The energy of the system is not constant, and we must take into account the work done by the electric field on the system. For example, considering just the particle in Figure 25.1*a* as our system, we obtain the energy diagram shown in **Figure 25.4**. The particle can have only kinetic energy, so its increase in kinetic energy is now due to work done by the electric field on it.

25.3 (*a*) Suppose the particle in Figure 25.4 moves along the x axis from point A at $x = x_A$ to point B at $x = x_B$. How much work is done by the electric field $\vec{E}$ on it? (*b*) Suppose now that an external agent moves the particle back from B to A, starting and ending at rest as shown in **Figure 25.5**. How much work does the electric field do on the particle as it is moved? (*c*) How much work does the agent do on the particle while it is moved? (*d*) What is the combined work done by the agent and by the electric field on the particle as it is moved? (*e*) Draw an energy diagram to illustrate the energy changes of the particle as the agent moves it from B to A.

Figure 25.5 Checkpoint 25.3.



CONCEPTS

Checkpoint 25.3 illustrates the importance of distinguishing between the work done by the electric field on a charged particle and the work done by the agent moving it. We shall refer to the work done by an electrostatic field as **electrostatic work.** If the particle begins and ends at rest, the electrostatic work is equal in magnitude and opposite in sign to the work done by the agent doing the moving. Then the total work done on the particle—the sum of the electrostatic and mechanical work—is zero. Indeed, the kinetic energy of the particle does not change, and because the particle possesses no other form of energy, its energy remains unchanged: $\Delta E = W = 0$.

Note that the electric force between charged particles, just like the gravitational force, is a central force: Its line of action always lies along the line connecting the two interacting particles (see Section 13.2). A direct consequence of this fact is:

The electrostatic work done on a charged particle as it moves from one point to another is independent of the path taken by the particle and depends on only the positions of the endpoints of the path.

The proof of this statement parallels the one for gravitational forces in Section 13.6. Imagine, for example, lifting a particle from A to B along curved path 2 in **Figure 25.6*a*** instead of along straight path 1. As shown in Figure 25.6*b*, the path can be approximated by small straight horizontal and vertical segments. Along the horizontal segments, the electric force is perpendicular to the force displacement, so the electrostatic work on the particle along these segments is zero. Along the vertical segments, the electrostatic work on the particle is nonzero, but note that each vertical segment corresponds to an equivalent vertical segment on path 1. Thus, the displacements along all the vertical segments of path 2 add up to precisely the displacement along path 1. In other words, the electrostatic work on the particle along path 2 (or any other path from A to B) is equal to that along path 1.

Figure 25.6*c* shows how this argument can be generalized to a nonuniform electric field. Imagine a particle being moved from point A to point C along the gray trajectory in the electric field caused by an object carrying a charge at the origin. We can approximate the trajectory by a succession of small circular arcs centered about the origin and small straight radial segments. The electrostatic work done on the particle along the circular arcs is zero because the force exerted on the particle is perpendicular to the force displacement. The radial segments, on the other hand, contribute to the electrostatic work done on the particle. The sum of all the radial segments, however, is equal to the radial displacement from A to B. Because no electrostatic work is done on the particle along the circular path from B to C, we thus see that the electrostatic work done on the particle along the gray trajectory from A to C is equal to the electrostatic work done along the path from A to B. The electrostatic work done on the particle along *any* path from A to C is thus the same as the electrostatic work done from A to B.

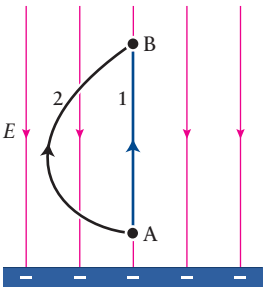
25.4 Suppose the electrostatic work done on a charged particle as it moves along the gray path from A to C in Figure 25.6*c* is W. What is the electrostatic work done on the particle (*a*) along the path from C to B to A and (*b*) along the closed path from A to C to B and back to A?

Checkpoint 25.4 shows that the electrostatic work done on a charged particle that moves around a closed path—*any* closed path—in an electrostatic field is zero.

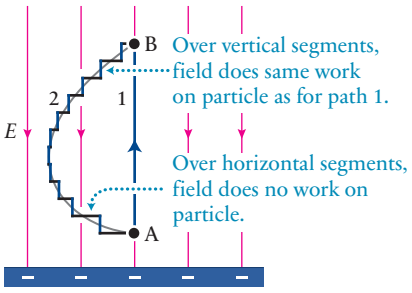
CONCEPTS

Figure 25.6 The electrostatic work done on a charged particle as the particle moves from point A to point B is independent of the path taken; it depends only on the positions of the endpoints of the path.

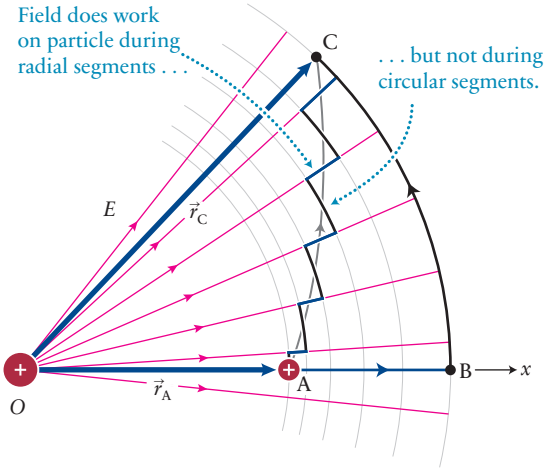
(*a*) Two paths by which particle can move from A to B



(*b*) Path 2 approximated by vertical & horizontal segments



(*c*) Same argument applied to nonuniform electric field



We obtained a similar result for the work done by the gravitational force (Eq. 13.17). The physical reason for this result is that the electric force between two charged particles, like the gravitational force between two objects that have mass, is nondissipative: When charged particles are moved around in electrostatic fields, no energy is irreversibly converted to other forms of energy. This is so because the electric interaction is nondissipative.

25.5 (*a*) If the electrostatic work done on a charged particle as it moves around a closed path is zero, does this also mean that the electrostatic work done on the particle is zero as it moves along a piece of the closed path? (*b*) In Checkpoint 25.3*a* we found that the electrostatic work done on the particle is the product of the x components of the electric force and the force displacement. Is the same true for the electrostatic work done on a particle that is moved along path AB in Figure 25.6*c*? (*c*) What happens to this electrostatic work if (*i*) the charge on the particle and (*ii*) the mass of the particle is doubled?

The third part of Checkpoint 25.5 demonstrates a very important point: The electrostatic work done on a charged particle is proportional to the charge carried by that particle. This means that once we have calculated the electrostatic work done on a particle carrying a charge q along some path in an electrostatic field, we don't need to carry out the whole calculation again if we are interested in the electrostatic work done on another particle carrying a charge $2q$. We know that the electrostatic work done on the second particle is twice the electrostatic work done on the first. Thus, if we know the electrostatic work done on a particle carrying a unit positive charge along some path, then we know the electrostatic work done on a particle carrying *any* charge between the same two points! We therefore introduce a new quantity, called **electrostatic potential difference** (or simply **potential difference**), defined as:

The potential difference between point A and point B in an electrostatic field is equal to the negative of the electrostatic work per unit charge done on a charged particle as it moves from A to B.

The potential difference is a scalar, and because the electrostatic work done on a charged particle as it moves from one position to another can be positive or negative, the potential difference can also be positive or negative; the potential difference between any two points B and A is the negative of that between points A and B. It is important to keep in mind that potential difference *is not a form of energy*—it is electrostatic work done per unit charge and therefore has SI units of J/C.

You may be wondering why the potential difference is defined in terms of the *negative* of the electrostatic work done on a particle and not in terms of the energy required to move the particle, which has the opposite sign. The reason is that the energy required to move the particle depends on the change in the particle's kinetic energy. If the particle starts at rest and ends at a nonzero speed, the particle's kinetic energy increases and so the energy required is greater than when it starts and ends at rest. The electrostatic work done on a particle, on the other hand, is independent of any change in the particle's kinetic energy.

25.6 (*a*) Is the potential difference along any path from A to C in Figure 25.6*c* positive, negative, or zero? (*b*) Along any path from C to B? (*c*) Along the straight path from B to A? (*d*) In Figure 25.4, is the potential difference between the particle's initial and final positions positive, negative, or zero? (*e*) Express this potential difference in terms of the change in the particle's kinetic energy ΔK and its charge q.

In principle, only the potential *difference* between the endpoints of a path is meaningful. We can, however, assign a value to the **potential** at each of these endpoints by choosing a reference point. Specifically, if there is a positive potential difference between points A and B, then A is at a lower potential than B; by assigning a value to the potential at one of the two points, the value of the potential at the other point is fixed. Potential and potential difference, which are immensely useful in solving problems in electrostatics, are discussed in more detail in Section 25.5.

25.3 Equipotentials

As we have seen, the electrostatic work done on a charged particle along the horizontal segments in Figure 25.6*b* and the circular arcs in Figure 25.6*c* is zero. Consequently, the potential difference between any two points on such an arc or horizontal segment is zero. In other words, the potential has the same value at all points along these arcs or segments. Such paths are said to be **equipotential lines:**

An equipotential line is a line along which the value of the electrostatic potential does not change. The electrostatic work done on a charged particle as it moves along an equipotential line is zero.

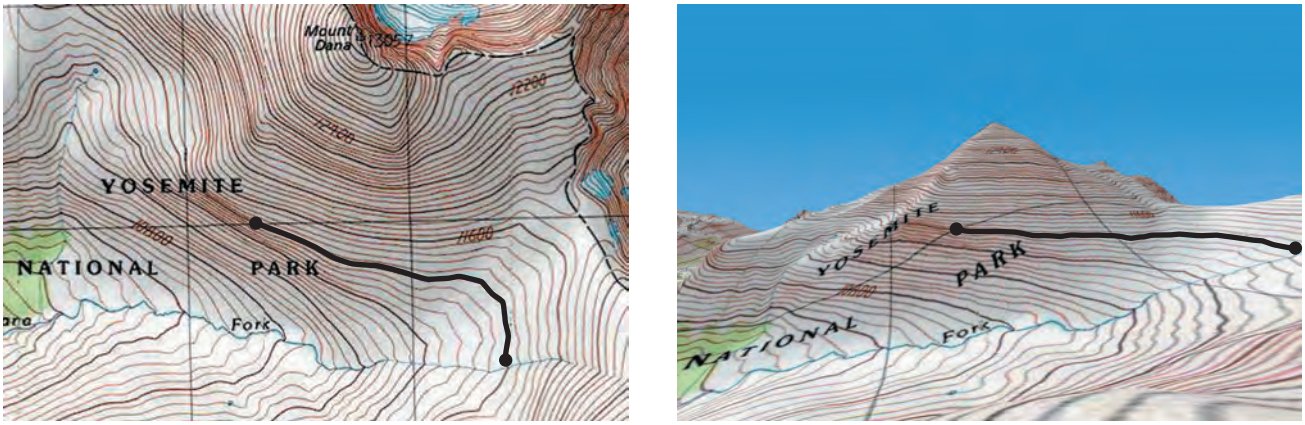
The equipotential lines in Figure 25.6 are much like contour lines on a topographical map. For example, the contour lines in **Figure 25.7** connect points of equal elevation. If you follow a contour line, your elevation remains the same, and so the gravitational potential energy associated with the separation between Earth and your body remains constant. Consequently, no gravitational work is done by Earth on you while you follow contour lines—these lines represent "gravitational equipotential lines."

Returning to the electrostatic case, note that Figure 25.6 is a cross section—a two-dimensional slice of a three-dimensional situation. Thus, the equipotential

CONCEPTS

Figure 25.7 Contour lines on a map are analogous to equipotential lines. If you hike along a contour line, shown in (*a*) two dimensions and (*b*) three dimensions, you neither gain nor lose gravitational potential energy.

(*a*) (*b*)

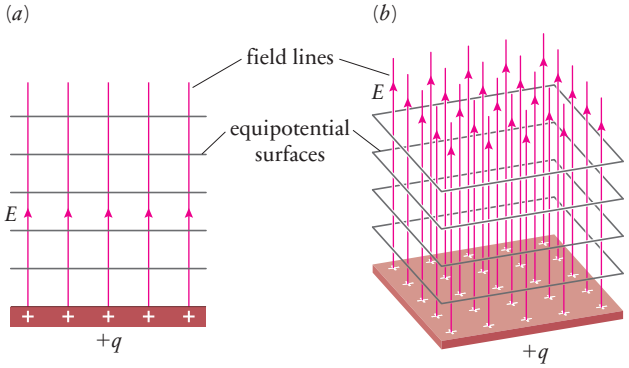


line segments shown are really parts of **equipotential surfaces.** In Figure 25.6*a*, for example, the electrostatic work done on a charged particle is zero along *any* displacement parallel to the surface of the negatively charged object, into or out of the plane of the drawing. Consequently, as shown in **Figure 25.8**, any surface parallel to the surface of the charged sheet causing the electric field is an equipotential surface. Often we'll use the term *equipotential* to denote an equipotential line or surface.

Just as with vector fields and field line diagrams, it is impossible to draw equipotential surfaces at all locations. In general, equipotentials are drawn with some fixed potential difference between them, just as contour lines represent a fixed difference in altitude. Keep in mind, however, that at any point between the equipotential surfaces shown in a figure we can, in principle, draw another equipotential surface.

25.7 Consider a single charged particle. Are there any equipotential lines or equipotential surfaces surrounding this particle?

Figure 25.8 Equipotential surfaces in a uniform electric field in (*a*) two dimensions; and (*b*) three dimensions.



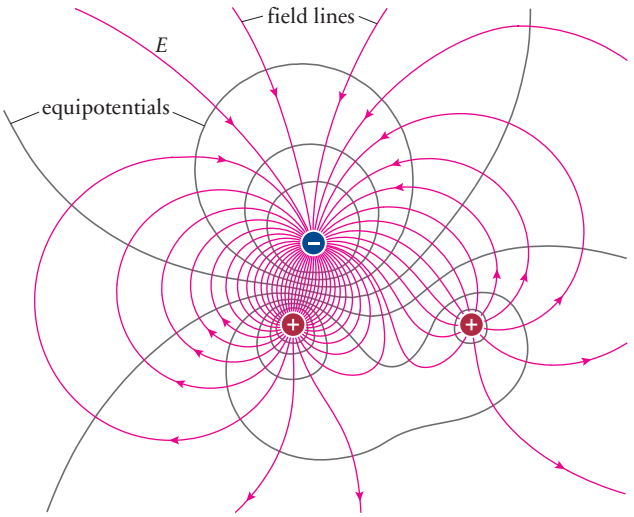
The equipotential surfaces in Figure 25.8 and those in Checkpoint 25.7 are perpendicular to the electric field lines. This is true for *any* stationary charge distribution:

The equipotential surfaces of a stationary charge distribution are everywhere perpendicular to the corresponding electric field lines.

The proof of this statement is straightforward: If the electric field line were not perpendicular to the equipotential surface, then the electric field would have a nonzero component along the surface. This means there would be a nonzero component of electric force along the surface. By definition, however, the electrostatic work done on a charged particle is zero along an equipotential surface, and so there cannot be such a component.

Figure 25.9 shows a two-dimensional view of the equipotential surfaces of a more complicated stationary charge distribution. Note how, at every point in the

Figure 25.9 Field lines and equipotentials for three stationary charged particles.



CONCEPTS

diagram, the equipotentials are, indeed, perpendicular to the field lines.

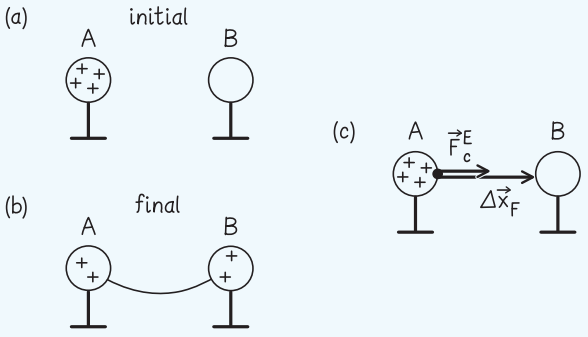
Recall from Section 24.5 that in electrostatic equilibrium, the electric field inside the bulk of a conducting object is zero, regardless of the shape of the object or any charge carried by it. This means that no electrostatic work is done on a charged particle inside a charged or uncharged conducting object. Thus, the entire volume of the conducting object is an *equipotential volume*. In electrostatic equilibrium, all points within a conducting object are at the same electrostatic potential.

EXAMPLE 25.1 Potential differences

Two metallic spheres A and B are placed on nonconducting stands. Sphere A carries a positive charge, and sphere B is electrically neutral. The two spheres are connected to each other via a wire, and the charge carriers reach a new electrostatic equilibrium. (*a*) Is the electric potential energy of the charge configuration after the spheres are connected greater than, smaller than, or equal to that of the original configuration? (*b*) Before the spheres are connected, is the potential difference between A and B positive, negative, or zero? Is it positive, negative, or zero after the spheres are connected?

1 GETTING STARTED I need to evaluate the electric potential energy—the energy associated with the configuration of charge carriers—and the potential difference—the negative of the electrostatic work per unit charge done on a charge carrier. I begin with two sketches: one showing the charge distribution before the spheres are connected, and one showing the distribution after they are connected (**Figures 25.10*a*** and *b*). Once the spheres are connected, they form one conducting object, and the positive charge initially on A spreads out over both spheres.

Figure 25.10



2 DEVISE PLAN To evaluate the potential difference between A and B, I need to determine the negative of the electrostatic work done on a charged particle as it is moved from A to B and then divide the result by the charge on the particle.

3 EXECUTE PLAN (*a*) After the two spheres are connected, the charge carriers spread out over both spheres, so they are farther apart than they were before the connection was made. Because I know that energy is required to push positively charged particles together, I conclude that the electric potential energy associated with the charge configuration in Figure 25.10*b* is smaller than that associated with the configuration in Figure 25.10*a*. ✔

(*b*) When a positive charge carrier is moved from A to B, the electrostatic work done on the carrier is *positive* because the electric force exerted on the carrier (directed away from sphere A) and the force displacement (from A to B) point in the same direction (Figure 25.10*c*). Consequently, the potential difference between A and B must be negative. ✔

Once the spheres are connected, they form one conducting object. Because the electric field is zero inside the entire object, no energy is required to move charge carriers around inside A and B, or from A to B, or vice versa. Thus, the potential difference between A and B after they are connected is zero. ✔

4 EVALUATE RESULT I know that a closed system always tends to arrange itself so as to lower the system's potential energy (see Section 7.8). It therefore makes sense that the electric potential energy of the system is smaller after the spheres are connected. That the potential difference between A and B is negative before they are connected is a direct consequence of the definition of potential difference. After A and B are connected, they form one large equipotential volume and so it makes sense that the potential difference is zero.

Example 25.1 allows us to make another very useful observation:

An electrostatic field is directed from points of higher potential to points of lower potential.

Because the electric field gives the direction of the force exerted on a positively charged particle, a direct consequence of this fact is:

In an electrostatic field, positively charged particles tend to accelerate toward regions of lower potential, whereas negatively charged particles tend to accelerate toward regions of higher potential.

CONCEPTS

25.8 When you hold a positively charged rod above a metallic sphere without touching it, a surplus of negative charge carriers accumulates at the top of the sphere, leaving a surplus of positive charge carriers at the bottom. Is the potential difference between the top and the bottom of the sphere positive, negative, or zero?

Self-quiz

1. Consider the situation illustrated in **Figure 25.11**. A positively charged particle is lifted against the uniform electric field of a negatively charged plate. Ignoring any gravitational interactions, draw energy diagrams for the following choices of systems: (*a*) particle and plate, $v_f = 0$; (*b*) particle only, $v_f = 0$; (*c*) particle, plate, and person lifting, $v_f = 0$; (*d*) particle and plate, $v_f \neq 0$.
2. A positively charged particle is moved from point A to point B in the electric field of the massive, stationary, positively charged object in **Figure 25.12**. (*a*) Is the electrostatic work done on the particle positive, negative, or zero? (*b*) How is the electrostatic work done on the particle along the straight path from A to B different from the electrostatic work done on the particle along a path from A to B via C?
3. **Figure 25.13** shows both the electric field lines and the equipotentials associated with the given charge distribution. (*a*) Is the potential at point A higher than, lower than, or the same as the potential at point B? (*b*) Is the potential at point C higher than, lower than, or the same as that at point B? (*c*) Is the potential at point C higher than, lower than, or the same as that at point A?

Figure 25.11

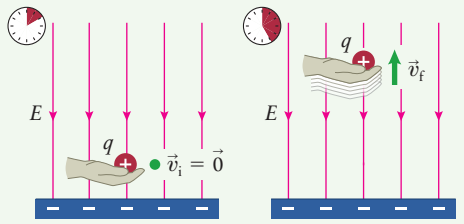


Figure 25.12

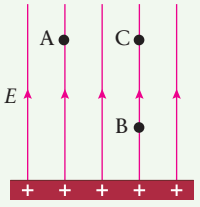
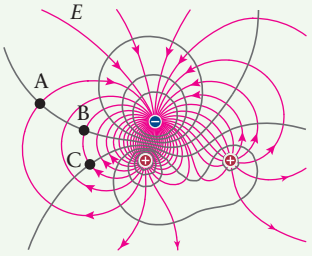


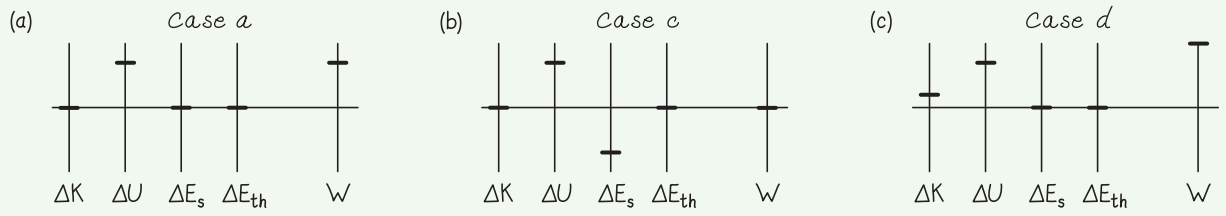
Figure 25.13



Answers

1. Only the person involves the conversion of source energy, so $\Delta E_s = 0$ in cases *a*, *b*, and *d*. When $v_f = 0$, $\Delta K = 0$; in *d*, $v_f \neq 0$ and so $\Delta K > 0$. For case *a*, the electric potential energy of the system increases as a result of positive work done by the agent (hand) on the particle, see **Figure 25.14a**. For case *b*, the (kinetic) energy of the particle alone does not change because the positive work done by the agent on the particle and the negative electrostatic work done on it are equal in magnitude, so all bars are zero. For case *c*, a decrease in the source energy (provided by the agent) is responsible for the increase in electric potential energy of the system (Figure 25.14*b*). For case *d*, the positive work done by the agent increases both the kinetic energy of the particle and the electric potential energy of the system (Figure 25.14*c*).

Figure 25.14



2. (*a*) The electrostatic work done on the particle while it moves from point A to point B is negative because the angle between the force exerted on the particle and the force displacement is between 90° and 180° (see Eqs. 10.35 and 10.33). (*b*) The electrostatic work done along the two paths is the same because the electrostatic work done on a particle between two points is independent of the path taken.
3. (*a*) The same, because these points lie along an equipotential surface. (*b*) Higher. The potential increases in a direction opposite to the direction of the electric field, so point C is at a higher potential than point B. (*c*) Higher. Points A and B are at the same potential, and point C is at a higher potential than point B.

25.4 Calculating work and energy in electrostatics

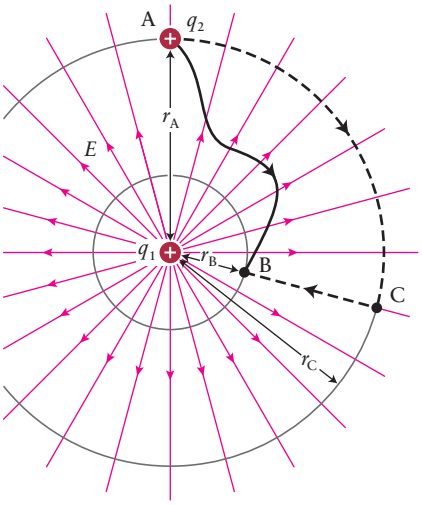


Figure 25.15 The electrostatic work done by particle 1 on particle 2 as the latter is moved from A to B is the same for the meandering solid path and for the dashed path ACB.

To quantify the electrostatic work done on a charged particle, consider charged particles 1 and 2 in **Figure 25.15**. Particle 2 is moved from point A to point B through the nonuniform electric field of particle 1, which is held stationary. The electrostatic work done by particle 1 on particle 2 as it is moved along the solid path from A to B is given by (Eq. 10.44)

$$W_{12}(\mathrm{A}\to\mathrm{B}) = \int_{\mathrm{A}}^{\mathrm{B}} \vec{F}_{12}^{E}\cdot d\vec{\ell}, \tag{25.1}$$

where $\vec{F}_{12}^{E}$ is the electric force exerted by particle 1 on particle 2 and $d\vec{\ell}$ is an infinitesimal vector segment of the path. This line integral (see Appendix B) is generally not easy to calculate because the magnitude of the electric force and the angle between the force and the path segment $d\vec{\ell}$ vary along the path. However, as we saw in Section 25.2, the electrostatic work done on particle 2 along the solid path from A to B in Figure 25.15 is the same as that done along the dashed path from A to C to B. The circular path from A to C is along an equipotential (see Checkpoint 25.7), so the electrostatic work done on the particle along that path is zero. The electric force is given by Coulomb's law (Eq. 22.7). If we take particle 1 to be at the origin, we can write $r_{12} = r$ for the distance between the two particles and $\hat{r}_{12} = \hat{r}$ for the unit vector pointing from particle 1 to particle 2. Along the radial path from C to B, an infinitesimal vector segment of the path can be written as $d\vec{\ell} = dr\,\hat{r}$, and so with $k = 1/(4\pi\epsilon_0)$ (Eq. 24.7), the integrand in Eq. 25.1 becomes

$$\vec{F}_{12}^{E}\cdot d\vec{\ell} = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2}\,\hat{r}\cdot(dr\,\hat{r}) = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2}\,dr. \tag{25.2}$$

The line integral in Eq. 25.1 thus becomes

$$W_{12}(\mathrm{A}\to\mathrm{B}) = W_{12}(\mathrm{C}\to\mathrm{B}) = \frac{q_1q_2}{4\pi\epsilon_0}\int_{r_C}^{r_B}\frac{1}{r^2}\,dr$$

$$= -\frac{q_1q_2}{4\pi\epsilon_0}\left[\frac{1}{r}\right]_{r_C}^{r_B} = \frac{q_1q_2}{4\pi\epsilon_0}\left[\frac{1}{r_C}-\frac{1}{r_B}\right]. \tag{25.3}$$

The distance from particle 1 to point A is the same as the distance from particle 1 to point C, so we have $r_A = r_C$ and

$$W_{12}(\mathrm{A}\to\mathrm{B}) = \frac{q_1q_2}{4\pi\epsilon_0}\left[\frac{1}{r_A}-\frac{1}{r_B}\right]. \tag{25.4}$$

Generalizing this result to arbitrary initial and final points, we get

$$W_{12} = \frac{q_1q_2}{4\pi\epsilon_0}\left[\frac{1}{r_{12,\mathrm{i}}}-\frac{1}{r_{12,\mathrm{f}}}\right], \tag{25.5}$$

where $r_{12,\mathrm{i}}$ and $r_{12,\mathrm{f}}$ are the initial and final values of the distance separating particles 1 and 2. Note that this expression is independent of the path taken: The electrostatic work done on particle 2 depends on only the distance between the two particles at the endpoints. Equation 25.5 also does not require particle 1 to be at the origin.

QUANTITATIVE TOOLS

25.9 (*a*) Using Eq. 25.5, determine whether the electrostatic work done on particle 2 along path CB in Figure 25.15 is positive, negative, or zero. (*b*) Does moving the particle carrying charge q_2 along path CB involve positive, negative, or zero mechanical work done on the particle? Assume the particle begins and ends at rest; verify the consistency of your answer with part *a*. (*c*) By how much does the electric potential energy of the two charged particles change as particle 2 is moved from C to B? Is this change positive, negative, or zero?

As Checkpoint 25.9 demonstrates, the change in electric potential energy of the system that comprises the two particles in Figure 25.15 is the negative of the electrostatic work done by particle 1 on the system that comprises particle 2 only. (This relationship between electrostatic work and change in electric potential energy can also be seen in Figures 25.1 and 25.4.) Taking the negative of Eq. 25.5, we thus obtain

$$\Delta U^E = \frac{q_1 q_2}{4\pi\epsilon_0}\left[\frac{1}{r_{12,\mathrm{f}}} - \frac{1}{r_{12,\mathrm{i}}}\right]. \tag{25.6}$$

Equation 25.6 gives only the *change* in electric potential energy, not *the* potential energy for a given configuration of charge. To obtain such an expression, we must first choose a zero of potential energy. It is customary to choose this zero to correspond to the configuration for which the force between the interacting particles is zero—that is to say, for infinite separation: $U^E = 0$ when $r_{12} = \infty$. Substituting this choice of reference point for the initial point in Eq. 25.6, we obtain

$$\Delta U^E \equiv U^E_{\mathrm{f}} - U^E_{\mathrm{i}} = U^E_{\mathrm{f}} - 0 = \frac{q_1 q_2}{4\pi\epsilon_0}\left[\frac{1}{r_{12,\mathrm{f}}} - 0\right]. \tag{25.7}$$

Thus, the **electric potential energy** for two particles carrying charges q_1 and q_2 and separated by distance r_{12} is

$$U^E = \frac{q_1 q_2}{4\pi\epsilon_0}\frac{1}{r_{12}} \quad (U^E \text{ zero at infinite separation}). \tag{25.8}$$

As this expression shows, the electric potential energy associated with two positively charged particles is positive: If they are brought together starting from infinite separation, the electric potential energy of the two particles increases. This is as it should be, because the two particles repel each other and so energy must be added to the system to bring it together ($W > 0$). The same idea applies to two negatively charged particles because the product $q_1 q_2$ is still positive. For a system of oppositely charged particles, on the other hand, $q_1 q_2 < 0$ and so $U^E < 0$. Indeed, the interaction is attractive and the potential energy decreases as the two are brought together.

25.10 Suppose we keep particle 2 stationary and move particle 1 in so that their final separation is the same as that in Figure 25.15. Is the electrostatic work done on particle 1 as it is moved in also given by the right-hand side of Eq. 25.4?

EXAMPLE 25.2 Putting two charged particles together

Two small pith balls, initially separated by a large distance, are each given a positive charge of 5.0 nC. By how much does the electric potential energy of the two-ball system change if the balls are brought together to a separation distance of 2.0 mm?

❶ **GETTING STARTED** I am asked to calculate a change in the electric potential energy of a system. I take *large* to mean a separation distance great enough that the balls don't interact. So the initial state is one in which the two pith balls have infinite separation. In the final state they are 2.0 mm apart.

(Continued)

QUANTITATIVE TOOLS

❷ DEVISE PLAN To calculate the change in electric potential energy I can use Eq. 25.6.

❸ EXECUTE PLAN Substituting $q_1 = q_2 = 5.0 \times 10^{-9}$ C, $r_{12,i} \approx \infty$, $r_{12,f} = 2.0 \times 10^{-3}$ m, and $k = 1/(4\pi\epsilon_0) = 9.0 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2$ into Eq. 25.6, I get

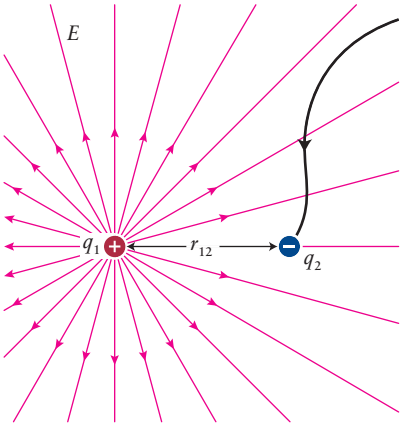
$$\Delta U^E = (9.0\times10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)(5.0\times10^{-9}\ \text{C})^2\left[\frac{1}{2.0\times10^{-3}\ \text{m}} - 0\right]$$

$$= 1.1 \times 10^{-4}\ \text{J}.$$

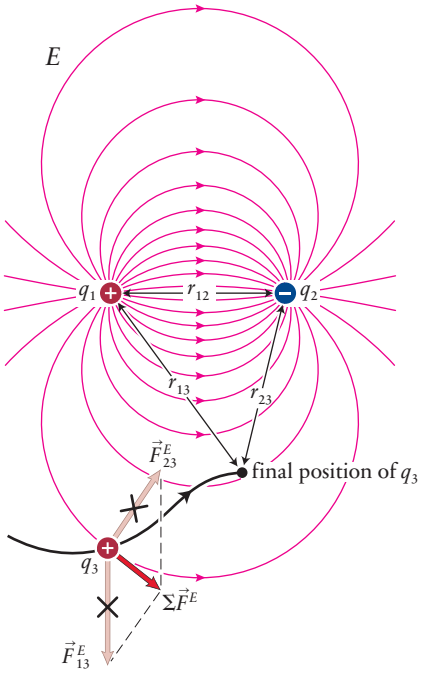
❹ EVALUATE RESULT My answer is positive, which makes sense because the balls repel each other and so work must be done on the system as they are brought together. This work increases the potential energy of the system. The magnitude of the potential energy change is small, but so is the magnitude of the force between the balls: Substituting q_1, q_2, and $r_{12,f}$ into Coulomb's law (Eq. 22.1), I obtain a force magnitude of 0.056 N. This is the maximum force between the two balls, so it makes sense that the energy associated with this interaction is small.

Figure 25.16 To obtain the electrostatic potential energy of a system of three charged particles, we assemble the system one particle at a time.

(*a*) We bring in second charged particle



(*b*) We bring in third charged particle



We can readily generalize the expressions for electrostatic work and electric potential energy for situations involving more than two charged particles. To determine the electric potential energy of a system of three charged particles in a certain configuration, for example, we calculate the electrostatic work done while assembling the system in its final configuration, starting from a situation where all three particles are far apart. Placing the first particle, carrying charge q_1, in its final position involves no electrostatic work because the other two particles are far away and therefore not interacting with particle 1. Next we bring in particle 2, carrying charge q_2, as illustrated in **Figure 25.16a**. The electrostatic work done while moving particle 2 can be found by substituting infinity for $r_{12,i}$ and r_{12} for $r_{12,f}$ in Eq. 25.5:

$$W_{12} = -\frac{q_1 q_2}{4\pi\epsilon_0}\frac{1}{r_{12}}. \tag{25.9}$$

Finally we bring in particle 3, carrying charge q_3, as shown in Figure 25.16*b*. The electrostatic work done while moving particle 3 is given by Eq. 25.1:

$$W_3 = \int_i^f (\Sigma\vec{F}^E_3)\cdot d\vec{\ell}, \tag{25.10}$$

where $\Sigma\vec{F}^E_3$ is the vector sum of the forces exerted on particle 3. Particle 3 is subject to two forces, one exerted by particle 1 and one by particle 2, so

$$\int_i^f (\Sigma\vec{F}^E_3)\cdot d\vec{\ell} = \int_i^f (\vec{F}^E_{13} + \vec{F}^E_{23})\cdot d\vec{\ell}$$

$$= \int_i^f \vec{F}^E_{13}\cdot d\vec{\ell} + \int_i^f \vec{F}^E_{23}\cdot d\vec{\ell} = W_{13} + W_{23}. \tag{25.11}$$

In other words, the electrostatic work done as 3 is moved to its final position is the sum of the electrostatic work done when only 1 is present plus that done when only 2 is present. Now we can apply Eq. 25.5 to each term in Eq. 25.11:

$$W_{13} + W_{23} = -\frac{q_1 q_3}{4\pi\epsilon_0}\frac{1}{r_{13}} - \frac{q_2 q_3}{4\pi\epsilon_0}\frac{1}{r_{23}}. \tag{25.12}$$

The total electrostatic work done on the three-particle system while assembling the charge configuration is the sum of Eqs. 25.9 and 25.12:

$$W = -\frac{q_1 q_2}{4\pi\epsilon_0}\frac{1}{r_{12}} - \frac{q_1 q_3}{4\pi\epsilon_0}\frac{1}{r_{13}} - \frac{q_2 q_3}{4\pi\epsilon_0}\frac{1}{r_{23}}. \tag{25.13}$$

QUANTITATIVE TOOLS

Alternatively, with our choice of zero at infinity, the electric potential energy of the system is

$$U^E = \frac{q_1 q_2}{4\pi\epsilon_0}\frac{1}{r_{12}} + \frac{q_1 q_3}{4\pi\epsilon_0}\frac{1}{r_{13}} + \frac{q_2 q_3}{4\pi\epsilon_0}\frac{1}{r_{23}} \quad (U^E \text{ zero at infinite separation}). \quad (25.14)$$

In words, to determine the electric potential energy of a system of charged particles, we need to add the electric potential energies for each pair of particles.

25.11 Suppose the charged particles in Figure 25.16 are assembled in a different order—say, 3 first, then 1, and finally 2. Do you obtain the same result as in Eq. 25.13 and Eq. 25.14?

25.5 Potential difference

The negative of the electrostatic work per unit charge done on a particle that carries a positive charge q from one point to another is defined as the **potential difference** between those points:

$$V_{AB} \equiv V_B - V_A \equiv \frac{-W_q(A \to B)}{q}. \quad (25.15)$$

Potential difference is a scalar, and the SI units of potential difference are joules per coulomb (J/C). In honor of Alessandro Volta (1745–1827), who developed the first battery, this derived unit is given the name **volt** (V):

$$1\ \text{V} \equiv 1\ \text{J/C}. \quad (25.16)$$

For example, if the electric field in **Figure 25.17** does −12 J of electrostatic work on a particle carrying charge $q_2 = +2.0$ C as it is moved from point A to point B (in other words, it requires +12 J of work by an external agent without the particle gaining any kinetic energy), then the potential difference V_{AB} between A and B is $-(-12\ \text{J})/(2.0\ \text{C}) = +6.0$ V. That is, the potential at B is 6.0 V higher than that at A.

Once we know the potential difference V_{AB} between A and B, we can obtain the electrostatic work done on *any* object carrying a charge q as it is moved along any path from A to B:

$$W_q(A \to B) = -qV_{AB}. \quad (25.17)$$

Keep in mind that the subscripts AB mean "from A to B." Because B is the final position, we write $V_{AB} \equiv V_B - V_A$. So when we refer to the "potential difference between A and B," we always mean the potential at B minus the potential at A.

Potential is important in practical applications because the potential difference between two points can be measured readily with a device called a *voltmeter* (we'll encounter these devices when we discuss electrical circuits in Chapters 31 and 32). In the next chapter we shall discuss the operation of another familiar device, the *battery*, which allows one to maintain a constant potential difference between two points. A 9-V battery, for example, maintains a +9-V-potential difference between its negative and positive terminals. (The positive terminal is at the higher potential, and thus the potential difference is *positive* when going from − to +.) For

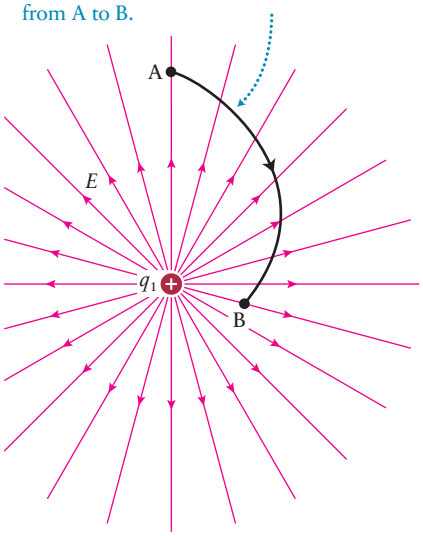


Figure 25.17 Once we know the electrostatic work done by a particle carrying a charge q_1 on a particle carrying a charge q_2 as the latter is moved from A to B in Figure 25.15, we can determine the potential difference between A and B in the electric field of particle 1.

example, when a particle carrying charge +1 C is moved from the negative terminal of a 9-V battery to the positive terminal, the particle undergoes a potential difference $V_{-+} = V_+ - V_- = +9$ V. Equation 25.17 tells us that the electrostatic work done on the particle is equal to

$$W_q(- \rightarrow +) = -qV_{-+} = -(+1\text{ C})(+9\text{ V}) = -9\text{ J}. \quad (25.18)$$

That this quantity is negative indicates that the agent moving the particle must do a positive amount of work on the particle. Likewise, when a particle carrying a charge of −2 C is moved from the − terminal to the + terminal of a 9-V battery, the electrostatic work done on the particle is $-(-2\text{ C})(+9\text{ V}) = +18$ J.

For simplicity, we shall denote the magnitude of the potential difference between the terminals of a battery by V_{batt}. Therefore

$$V_{\text{batt}} = V_+ - V_-. \quad (25.19)$$

Just as with potential energy, only potential *differences* are physically relevant. If we choose a reference point, however, we can determine the value for the potential at any other point. In the preceding section we chose infinity as the reference point for the electric potential energy for charged particles because they do not interact at infinite distance: $U^E(\infty) = 0$. The same choice can be made for the potential, but when we deal with electrical circuits it is customary to assign zero potential to Earth (ground) because Earth is a good and very massive conducting object through which the motion of charge carriers requires negligible energy.

EXERCISE 25.3 Potential and potential difference

The negative terminal of a 9-V battery is connected to ground via a wire. (*a*) What is the potential of the negative terminal? (*b*) What is the potential of the positive terminal? (*c*) What is the potential of the negative terminal if the positive terminal is connected to ground?

SOLUTION (*a*) I know from Section 25.3 that any conducting objects that are in electrical contact with each other form an equipotential. Once they are in contact, therefore, the ground (which is conducting), the wire, and the negative terminal are all at the same potential. If the potential of the ground is zero (an arbitrary but customary choice), then the potential of the negative terminal is also zero. ✔

(*b*) The potential difference between the negative and positive terminals is +9 V, meaning that the potential of the positive terminal is 9 V higher than that of the negative terminal: $V_{\text{batt}} = V_+ - V_- = +9$ V. If V_- is zero, then the potential of the positive terminal is $V_+ = +9\text{ V} + V_- = (+9\text{ V}) + (0\text{ V}) = +9$ V. ✔

(*c*) With the positive terminal connected to ground, that terminal's potential becomes zero. Because the battery maintains a potential difference of +9 V between the negative and positive terminals, I now have $V_{\text{batt}} = V_+ - V_- = 0 - V_- = +9$ V and so the negative terminal is at a potential of –9 V. ✔

QUANTITATIVE TOOLS

To obtain an explicit expression for the potential difference between two points A and B in the electric field of particle 1 carrying charge q_1, we start with the expression for the electrostatic work done by particle 1 on a particle 2 carrying charge q_2 as it is moved from A to B (Eq. 25.4). All we need to do is add a minus sign and divide by q_2:

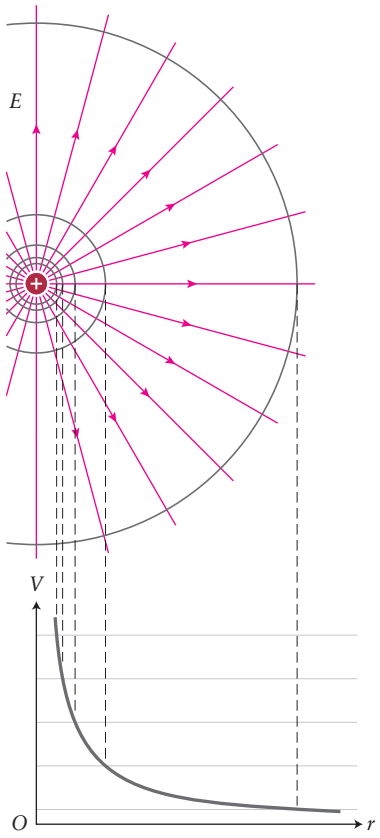
$$V_{\text{AB}} \equiv \frac{-W_{12}(\text{A} \rightarrow \text{B})}{q_2} = \frac{q_1}{4\pi\epsilon_0}\left[\frac{1}{r_\text{B}} - \frac{1}{r_\text{A}}\right]. \quad (25.20)$$

For charged particles, the choice of ground as a zero for the potential is not very meaningful. However, if we set the zero for potential at infinity and let r_A be at infinity, we obtain for the potential at a distance $r = r_B$ from a single charged particle

$$V(r) = \frac{1}{4\pi\epsilon_0}\frac{q_1}{r} \quad \text{(potential zero at infinity)}. \tag{25.21}$$

Equation 25.21 confirms that the potential is constant on any spherical surface centered on the charged particle (that is, any surface for which r is constant), as we concluded in Checkpoint 25.7. **Figure 25.18** shows the potential for a charged particle in graphical form. The bottom of the figure shows the $1/r$ dependence of the potential; the top shows a two-dimensional view of the equipotentials around the charged particle. The $1/r$ dependence of the potential is reflected by the increasing spacing of the equipotentials (with a fixed potential difference between them) as r increases.

Figure 25.18 Equipotentials, field lines, and graph of potential for a charged particle.



25.12 (*a*) Using Eq. 25.20, determine whether the potential difference between A and B in Figure 25.17 is positive, negative, or zero. (*b*) From the directions of the electric force and the force displacement, is the electrostatic work done on a positively charged particle as it is moved along a straight path from A to B positive, negative, or zero? Verify that your answer is consistent with the answer in part *a*.

EXAMPLE 25.4 Atomic potential difference

A (simplistic) model of the hydrogen atom treats the electron as a particle carrying a charge $-e$ orbiting a proton (a particle carrying a charge $+e$) in a circle of radius $r_H = 0.53 \times 10^{-10}$ m. (*a*) How much energy is required to completely separate the electron from the proton? For simplicity, ignore the electron's kinetic energy. (*b*) Across what potential difference does the electron travel as it is separated from the proton?

❶ **GETTING STARTED** To completely separate the electron from the proton, I must increase the distance between them from r_H to infinity (at which point their interaction is reduced to zero). By increasing the separation, I increase the potential energy of the electron-proton system. The energy I must add to the system in order to separate the particles is equal to the increase in potential energy of the system.

❷ **DEVISE PLAN** I can calculate the potential energy increase from Eq. 25.6, setting $r_i = r_H$ and $r_f = \infty$. I can obtain the potential difference between the initial and final positions of the electron by taking the negative of the electrostatic work done by the electric field of the proton on the electron during the separation and dividing this value by the charge on the electron.

❸ **EXECUTE PLAN** (*a*) The change in the electric potential energy of the two-particle system is

$$\Delta U^E = \frac{q_1 q_2}{4\pi\epsilon_0}\left[\frac{1}{r_f} - \frac{1}{r_i}\right] = \frac{-e^2}{4\pi\epsilon_0}\left[0 - \frac{1}{r_H}\right]$$

$$= (9.0\times10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)\frac{(1.6\times10^{-19}\ \text{C})^2}{0.53\times10^{-10}\ \text{m}} = 4.3\times10^{-18}\ \text{J}.$$

Because there are no changes in any other forms of energy in the system, the energy required to separate the electron and the proton in a hydrogen atom is 4.3×10^{-18} J. ✔

(*b*) The answer I obtained in part *a* is the mechanical work an external agent must do on the electron to separate it from the proton. I know that the energy of the electron does not change because the electron gains no kinetic energy. Considering just the electron as my system, I thus have $\Delta E = 0$ and so the work done on the electron is zero. This work has two parts: mechanical work done by the external agent and electrostatic work done by the electric field of the proton. Because the sum of these two parts is zero, I know that the electrostatic work done by the electric field of the proton on the electron must be the negative of the mechanical work done by the external agent, so $W_{pe}(r_H \to \infty) = -4.3 \times 10^{-18}$ J. The potential difference is thus (Eq. 25.15)

$$V_{H\infty} = \frac{-W_{pe}(r_H \to \infty)}{q} = \frac{-(-4.3\times10^{-18}\ \text{J})}{-1.6\times10^{-19}\ \text{C}} = -27\ \text{V}. \checkmark$$

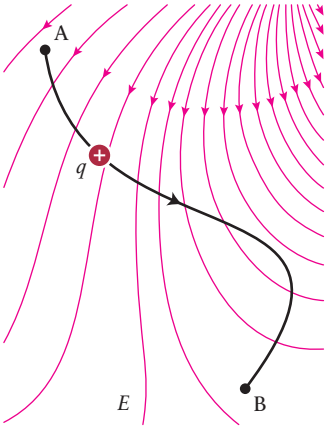
❹ **EVALUATE RESULT** The positive sign on ΔU^E in part *a* means that the potential energy of the proton-electron system increases when the two are moved apart. Therefore mechanical work must be done to pull them apart, as I expect because the electron and proton attract each other. The value I obtain is extremely small, but I know that 1 m^3 of matter contains about 10^{29} atoms (see Exercise 1.6), so

(Continued)

the electric potential energy in a cubic meter of matter is on the order of $(10^{28})(4.3 \times 10^{-18}\ \text{J}) \approx 10^{11}\ \text{J}$. In the box "Coherent versus incoherent energy" on page 221 I learned that the amount of chemical energy in a pencil is on the order of 10^5 J. Given that chemical energy is derived from electric potential energy stored in chemical bonds and that the volume of a pencil is about $10^{-5}\ \text{m}^3$, my answer for part *a* is not unreasonable.

For part *b*, I could have used Eq. 25.20 directly. Substituting the values given into that equation, I obtain the same answer, which gives me confidence in the answer I obtained.

Figure 25.19 To determine the potential difference between two points in an electric field, we must evaluate the electrostatic work done on a charged particle as the particle is moved along a path between those points.



To obtain a more general result for the potential difference between one point and another in an arbitrary electric field $\vec{E}$, consider the situation illustrated in **Figure 25.19**. A particle carrying a charge q is moved from point A to point B in an electric field due to some charge distribution (not visible in the illustration). The electrostatic work done on the particle is

$$W_q(\text{A} \to \text{B}) = \int_{\text{A}}^{\text{B}} \vec{F}_q^E \cdot d\vec{\ell}. \tag{25.22}$$

The vector sum of the forces exerted on the particle is equal to the product of the electric field and the charge q (Eq. 23.6):

$$\vec{F}_q^E = q\vec{E}, \tag{25.23}$$

so

$$W_q(\text{A} \to \text{B}) = q\int_{\text{A}}^{\text{B}} \vec{E} \cdot d\vec{\ell}. \tag{25.24}$$

The potential difference between point A and point B is therefore

$$V_{\text{AB}} \equiv \frac{-W_q(\text{A} \to \text{B})}{q} = -\int_{\text{A}}^{\text{B}} \vec{E} \cdot d\vec{\ell}. \tag{25.25}$$

In evaluating this line integral, we must keep in mind that the integral does not depend on the path taken but only on the endpoints. It therefore pays to choose a path that facilitates evaluating the integral. The Procedure box below and the next example will help you gain practice calculating potential differences between two points.

PROCEDURE: Calculating the potential difference between two points in an electric field

The potential difference between two points in an electric field is given by Eq. 25.25. The following steps will help you evaluate the integral:

1. Begin by making a sketch of the electric field, indicating the points corresponding to the two points between which you wish to determine the potential difference.
2. To facilitate evaluating the scalar product $\vec{E} \cdot d\vec{\ell}$, choose a path between the two points so that $\vec{E}$ is either parallel or perpendicular to the path. If necessary, break the path into segments. If $\vec{E}$ has a constant value along the path (or a segment of the path), you can pull it out of the integral; the remaining integral is then equal to the length of the corresponding path (or the segment of the path).
3. Remember that to determine V_{AB} ("the potential difference between points A and B"), your path begins at A and ends at B. The vector $d\vec{\ell}$ therefore is tangent to the path, in the direction that leads from A to B (see also Appendix B).

At this point you can substitute the expression for the electric field and carry out the integral. Once you are done, you may want to verify the algebraic sign of the result you obtained: *negative* when a positively charged particle moves along the path in the direction of the electric field, and *positive* when it moves in the opposite direction.

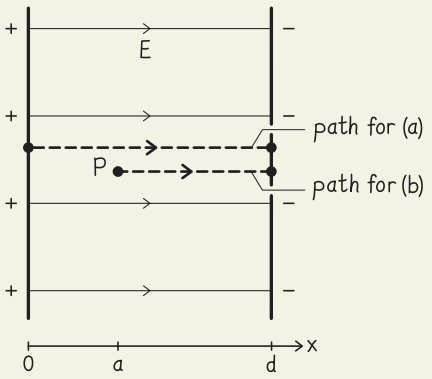
QUANTITATIVE TOOLS

EXAMPLE 25.5 Electrostatic potential in a uniform field

Consider a uniform electric field of magnitude E between two parallel charged plates separated by a distance d. (*a*) What is the potential difference between the positive plate and the negative plate? (*b*) What is the value of the potential at a point P that lies between the plates and is a distance $a < d$ from the positive plate, if the potential of the negative plate is zero?

❶ **GETTING STARTED** I begin by making a sketch of the two parallel plates and the electric field (**Figure 25.20**). For part *a* I must calculate the potential difference between the positive and negative plates, so I choose a path that begins at the positive plate and ends at the negative plate. For part *b* I am asked to determine, relative to the potential at the negative plate, the potential at a point P located between the plates a distance a from the positive plate, so I choose a path that runs from P to the negative plate. I indicate the endpoints for both paths in my drawing and choose my x axis to be parallel with the electric field with the positive plate at $x = 0$.

Figure 25.20



❷ **DEVISE PLAN** To obtain the potential difference between the endpoints of each path, I apply Eq. 25.25. With my choice of $+x$ axis the $d\vec{\ell}$ vector in Eq. 25.25 becomes $d\vec{\ell} = dx\,\hat{\imath}$. Because $\vec{E} = E\,\hat{\imath}$ I thus have $\vec{E}\cdot d\vec{\ell} = E\,dx$, and because the field is uniform, E is constant. For part *b*, once I have the potential difference between the endpoints of the path from P to the negative plate, I can determine the potential at P because I am told that the potential at the negative plate is zero.

❸ **EXECUTE PLAN** (*a*) Applying Eq. 25.25 to the path from $x_i = 0$ (positive plate) to $x_f = d$ (negative plate), I get

$$V_{0d} = -\int_0^d E\,dx = -E\int_0^d dx = -Ed, \checkmark \qquad (1)$$

where I have pulled E out of the integral because it is constant.

(*b*) If I start the line integral in Eq. 1 at point P, I obtain

$$V(d) - V(a) = -E\int_a^d dx = -E(d-a)$$

or, because $V(d) = 0$,

$$V(a) = E(d-a). \checkmark$$

❹ **EVALUATE RESULT** The negative sign in Eq. 1 means that the potential at the end of the path (that is, at the negative plate) is *lower* than the potential at the beginning (the positive plate). This negative potential difference is in agreement with the sign of the electrostatic work done on a positively charged particle: Positive electrostatic work is done on the particle as it is moved from the positive plate to the negative plate, and so, according to Eq. 25.15, the potential difference should indeed be negative.

My result for part *b* indicates that the potential is positive at $x = a$ $(a < d)$, decreases linearly with the distance a to the positive plate, and goes to zero at $a = d$. That makes sense, because I know that the potential of the positive plate must be higher than that of the negative plate and if $a = d$, I get $V(d) = 0$, as expected.

With an appropriate choice of reference point for the potential, Eq. 25.25 allows us to assign values to the potential at every point surrounding a charge distribution. This "potential field" is related to the electric field: Each can be determined from the other. A drawing that shows a set of equipotentials for a charge distribution is equivalent to a drawing that shows a set of field lines for that charge distribution. In Section 25.7 we shall show how the electric field can be derived from the potential field. The potential field, however, has advantages over the electric field. First, it is a scalar field whereas the electric field is a vector field, and so calculations involving the potential are generally simpler. Second, while no devices exist to measure electric field strength directly, we can measure the potential difference between two points with a voltmeter.

25.13 Verify that Eq. 25.25 is consistent with Eq. 25.20 by substituting the expression for the electric field of a charged particle.

QUANTITATIVE TOOLS

By following the same procedure as in Checkpoint 25.13, we can now obtain the potential for a group of charged particles. Recall that the electric field due to a group of charged particles is equal to the sum of the electric fields of the individual charged particles (Eq. 23.5),

$$\vec{E} = \sum_n \vec{E}_n. \tag{25.26}$$

This gives us

$$V_{AB} = \int_A^B \vec{E} \cdot d\vec{\ell} = -\int_A^B \left(\sum_n \vec{E}_n \right) \cdot d\vec{\ell}. \tag{25.27}$$

Because the integral of a sum is equal to a sum of integrals, we have

$$-\int_A^B \left(\sum_n \vec{E}_n \right) \cdot d\vec{\ell} = -\sum_n \int_A^B \vec{E}_n \cdot d\vec{\ell}. \tag{25.28}$$

The line integral after the summation sign is the negative of the potential difference due to the field $\vec{E}_n$, so we see that the total potential difference is the sum of the potential differences caused by the individual particles:

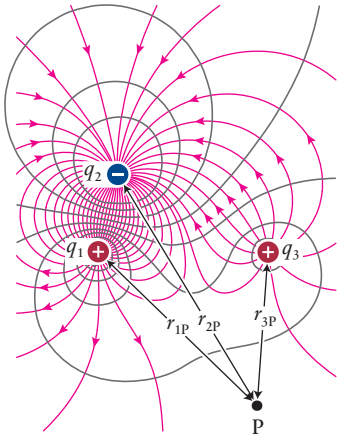
$$V_{AB} = \sum_n V_{AB,n}, \tag{25.29}$$

where $V_{AB,n}$ is the potential difference caused by particle n. Substituting the potential of a single particle, Eq. 25.21, and again letting the potential at infinity be zero, we get

$$V_P = \frac{1}{4\pi\epsilon_0} \sum_n \frac{q_n}{r_{nP}} \quad \text{(potential zero at infinity)}, \tag{25.30}$$

where q_n is the charge carried by particle n and r_{nP} is the distance of particle n from the point P at which we are evaluating the potential (**Figure 25.21**).

Figure 25.21 The potential at point P due to a group of charged particles is the algebraic sum of the potentials at P due to the individual particles.



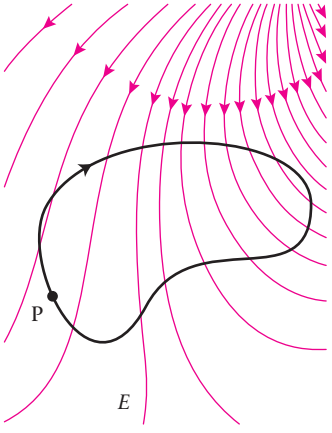
The line integral on the right in Eq. 25.25 has an important significance. As we argued in Section 25.2, the electrostatic work done on a charged particle moving around a closed path (as in **Figure 25.22**) is zero. Therefore, Eq. 25.24 must yield zero for a closed path:

$$W_q(\text{closed path}) = q \oint \vec{E} \cdot d\vec{\ell} = 0, \tag{25.31}$$

where the circle through the integral sign indicates that the integration is to be taken around a closed path. Because Eq. 25.31 holds for any value of q, we have

$$\oint \vec{E} \cdot d\vec{\ell} = 0 \quad \text{(electrostatic field)}. \tag{25.32}$$

Figure 25.22 The electrostatic work done on a charged particle as the particle is moved around a closed path starting and ending at some point P is zero.



In other words, for any electrostatic field, the line integral of the electric field around a closed path is zero. This is equivalent to saying that you cannot extract energy from an electrostatic field by moving a charged particle around a closed path. In terms of potential, this means that if we start at some point P on the closed path and the potential has a value V_P at that point, then the potential can take on other values as we go around the closed path, but as we return to P, the value of the potential must once again be V_P. As we shall see later in Chapter 30, it *is* possible to get energy out of electric fields, however. Here we have shown only that it is *not* possible to extract energy by moving a charged particle around a

closed path in *electrostatic* fields—that is, those due to stationary charge distributions.

25.14 Describe how the potential varies as you go around the closed path in Figure 25.22 in the direction shown, starting from a potential V_P at P.

25.6 Electrostatic potentials of continuous charge distributions

For extended objects with continuous charge distributions, we cannot use Eq. 25.30 directly to calculate the potential. Instead we must divide the object into infinitesimally small segments, each carrying charge dq_s (which we can treat as a charged particle), and then integrate over the entire object.

Consider, for example, the object shown in **Figure 25.23**. Let the zero of potential again be at infinity. Treating each segment as a charged particle, we calculate its contribution to the potential at P (Eq. 25.21):

$$dV_s = \frac{1}{4\pi\epsilon_0}\frac{dq_s}{r_{sP}}, \tag{25.33}$$

where r_{sP} is the distance between P and dq_s. The potential due to the entire object is then given by the sum over all the segments that make up the object. For infinitesimally small segments, this yields the integral

$$V_P = \int dV_s = \frac{1}{4\pi\epsilon_0}\int\frac{dq_s}{r_{sP}} \quad \text{(potential zero at infinity)}, \tag{25.34}$$

where the integral is taken over the entire object. Note the parallel between this integral and the integral in Eq. 23.15 for calculating the electric field of a continuous charge distribution. Indeed, the procedure for calculating the potential of a continuous charge distribution is very similar to that for calculating the electric field of a continuous charge distribution (see Section 23.7). However, the potential in Eq. 25.34 is much easier to evaluate because it is a scalar and does not involve any unit vectors. The Procedure box below provides some helpful steps in evaluating Eq. 25.34, and the next two examples show how to put the procedure into practice.

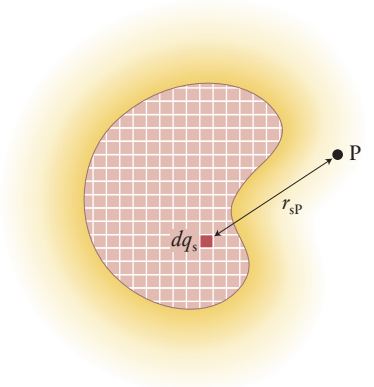


Figure 25.23 The potential due to an extended object is the algebraic sum of the potentials of all the infinitesimally small segments that make up the object.

PROCEDURE: Calculating the electrostatic potentials of continuous charge distributions

To calculate the electrostatic potential of a continuous charge distribution (relative to zero potential at infinity), you need to evaluate the integral in Eq. 25.34. The following steps will help you work out the integral:

1. Begin by making a sketch of the charge distribution. Mentally divide the distribution into infinitesimally small segments, each carrying a charge dq_s. Indicate one such segment in your drawing.
2. Choose a coordinate system that allows you to express the position of dq_s in the charge distribution in terms of a minimum number of coordinates (x, y, z, r, or θ). These coordinates are the integration variables. For example, use a radial coordinate system for a charge distribution with radial symmetry. Never place the representative segment dq_s at the origin.
3. Indicate the point at which you wish to determine the potential. Express the factor $1/r_{sP}$, where r_{sP} is the distance between dq_s and the point of interest, in terms of the integration variable(s).
4. Determine whether the charge distribution is one dimensional (a straight or curved wire), two dimensional (a flat or curved surface), or three dimensional (any bulk object). Express dq_s in terms of the corresponding charge density of the object and the integration variable(s).

At this point you can substitute your expressions for dq_s and $1/r_{sP}$ into Eq. 25.34 and work out the integral.

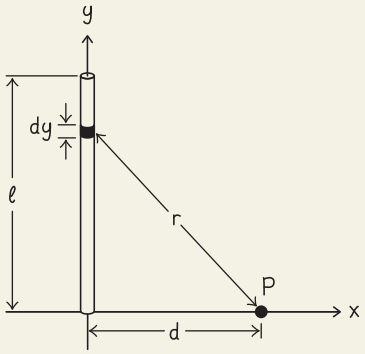
QUANTITATIVE TOOLS

EXAMPLE 25.6 Electrostatic potential of a uniformly charged thin rod

A thin rod of length ℓ carries a uniformly distributed charge q. What is the potential V_P at point P a distance d from the rod along a line that runs perpendicular to the long axis of the rod and passes through one end of the rod?

1 GETTING STARTED I begin by making a sketch. I let the y axis be along the rod, with the origin at the bottom of the rod. Point P then lies on the x axis (**Figure 25.24**). I take the potential at infinity to be zero. What I must calculate is the electrostatic potential V_P at P.

Figure 25.24



2 DEVISE PLAN Because the rod is thin and uniformly charged, I can treat it as a one-dimensional object that has a linear charge density $\lambda = q/\ell$. To determine the electrostatic potential at point P, I divide the rod lengthwise into a large number of infinitesimally small segments, each of length dy. Each segment contributes to the potential at P an amount given by Eq. 25.33. To calculate the potential at P due to the entire rod, I can then use Eq. 25.34 to integrate my result over the entire length of the rod.

3 EXECUTE PLAN Each length segment dy carries a charge $dq_s = \lambda dy = (q/\ell)\,dy$. To calculate the potential at P due to the whole rod, I substitute the distance between each segment dy and P, $r_{sP} = \sqrt{y^2 + d^2}$, and my expression for dq_s into Eq. 25.34 and integrate from the bottom of the rod at $y = 0$ to the top at $y = \ell$:

$$V_P = \frac{1}{4\pi\epsilon_0}\frac{q}{\ell}\int_0^{\ell}\frac{dy}{\sqrt{y^2 + d^2}}.$$

Looking up the solution of the integral, I obtain

$$V_P = \frac{1}{4\pi\epsilon_0}\frac{q}{\ell}\left[\ln(y + \sqrt{y^2 + d^2})\right]_0^{\ell}$$

$$= \frac{1}{4\pi\epsilon_0}\frac{q}{\ell}\left[\ln(\ell + \sqrt{\ell^2 + d^2}) - \ln d\right]$$

$$= \frac{1}{4\pi\epsilon_0}\frac{q}{\ell}\ln\left(\frac{\ell + \sqrt{\ell^2 + d^2}}{d}\right). \checkmark$$

4 EVALUATE RESULT: If I let ℓ go to zero, my answer should become the same as that for a particle (Eq. 25.21). When $\ell \ll d$, I can ignore the term ℓ^2 in my answer and so the argument of the logarithm becomes $(\ell + d)/d = 1 + \ell/d$. Because $\ln(1 + \epsilon) \approx \epsilon$ when $\epsilon \ll 1$ (see Appendix B), my answer becomes

$$V_P \approx \frac{1}{4\pi\epsilon_0}\frac{q}{\ell}\ln\left(\frac{\ell + d}{d}\right) \approx \frac{1}{4\pi\epsilon_0}\frac{q}{\ell}\frac{\ell}{d} = \frac{1}{4\pi\epsilon_0}\frac{q}{d},$$

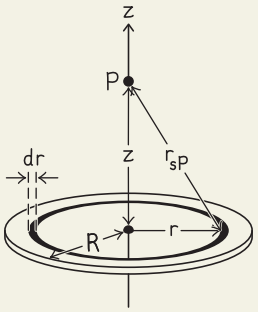
which is indeed equal to the electrostatic potential at a distance $r = d$ from a particle.

EXAMPLE 25.7 Electrostatic potential of a uniformly charged disk

A thin disk of radius R carries a uniformly distributed charge. The surface charge density on the disk is σ. What is the electrostatic potential due to the disk at point P that lies a distance z from the plane of the disk along an axis that runs through the disk center and is perpendicular to the plane of the disk?

1 GETTING STARTED I begin by making a sketch of the disk (**Figure 25.25**). I let the disk be in the xy plane, with the origin at the center of the disk and point P on the z axis.

Figure 25.25



2 DEVISE PLAN Because of the circular symmetry of the disk, I divide it into a large number of thin circular ring segments, each of radius r and thickness dr. All parts of a given ring are the same distance r_{sP} from point P, so each part makes the same contribution to the potential at P. I can therefore calculate the contribution of an entire ring segment to the potential at P using Eq. 25.33, substituting the charge dq_s on the segment and the distance from a point on the segment to P. The charge dq_s on the ring segment is given by the product of its area (circumference times thickness) and the surface charge density: $dq_s = (2\pi r)\,dr\,\sigma$. To calculate the potential at point P due to the entire disk, I can use Eq. 25.34 to integrate my result over the radius of the disk, using r as my integration variable.

3 EXECUTE PLAN: The distance from a point on any ring segment to P is $r_{sP} = \sqrt{z^2 + r^2}$, and so each segment's contribution to the potential is given by

$$dV_P = \frac{1}{4\pi\epsilon_0}\frac{(2\pi r)\sigma\,dr}{\sqrt{z^2 + r^2}}.$$

To calculate the potential at P due to the whole disk, I integrate this expression from $r = 0$ to $r = R$:

$$V_P = \int dV_P = \frac{1}{4\pi\epsilon_0}\int_0^R \frac{2\pi r\sigma}{\sqrt{z^2+r^2}}\,dr.$$

Looking up the solution of the integral, I get

$$V_P = \frac{2\pi\sigma}{4\pi\epsilon_0}\int_0^R \frac{r\,dr}{\sqrt{z^2+r^2}} = \frac{\sigma}{2\epsilon_0}\left(\sqrt{z^2+R^2}-|z|\right). \checkmark$$

4 EVALUATE RESULT When z is very large relative to R, the disk should resemble a particle and my result should reduce to that for a particle (Eq. 25.21). For large $z > 0$ I can use the binomial expansion (see Appendix B) to write the factor in parentheses as

$$\sqrt{z^2(1+R^2/z^2)} - z \approx z\left(1+\tfrac{1}{2}\frac{R^2}{z^2}\right) - z = \tfrac{1}{2}\frac{R^2}{z}.$$

Because $\sigma\pi R^2$ is equal to the charge q on the disk, my expression for V becomes

$$V_P \approx \frac{\sigma}{2\epsilon_0}\left(\tfrac{1}{2}\frac{R^2}{z}\right) = \frac{\sigma\pi R^2}{4\pi\epsilon_0 z} = \frac{1}{4\pi\epsilon_0}\frac{q}{z},$$

which is equal to the result for the potential along a z axis due to a particle located at the origin.

You may have noticed that the two examples above are parallel to corresponding examples in Section 23.7 in which we calculated the electric fields due to a thin charged rod or disk. If you compare the calculations, however, the advantage of working with the potential becomes obvious: Because the potential is a scalar you don't need to take vector sums. Haven't we thrown away some information, though? After all, the answer we get is also just a scalar, not a vector like the electric field. Figure 25.9 gives us some idea of the answer to this question: Because field lines and equipotentials are always perpendicular to each other, you can draw equipotentials if you know the field line pattern or, conversely, draw field lines if you know the equipotentials. It turns out that even though the potential is a scalar and the electric field is a vector, it is possible to determine one from the other.

25.15 Verify that the potentials obtained in Examples 25.6 and 25.7 have the correct sign.

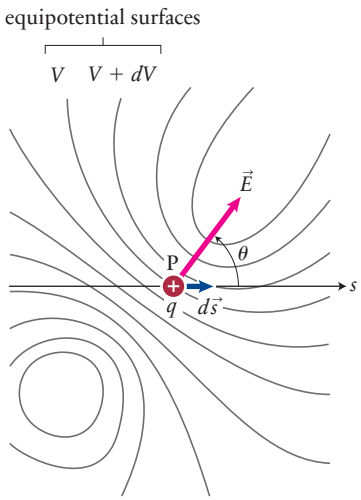
25.7 Obtaining the electric field from the potential

For the potential to be a useful quantity, we must be able to determine the electric field (and therefore the forces exerted by this field) from the potential. For example, let the equipotentials in **Figure 25.26** represent the potential of some charge distribution. How can we use the known potential of the charge distribution to determine the value of the electric field at any point P?

We know that the electric field is perpendicular to the equipotentials, and so the electric field at P must be along the direction indicated in the figure. To determine the magnitude of the electric field, imagine moving a particle carrying a charge q over an infinitesimally small displacement $d\vec{s}$ along some arbitrary axis s. Let the particle be displaced from P, where the potential is V, to a point P′ where the potential is $V + dV$. According to Eq. 25.17, the electrostatic work done on the particle is

$$W_q(\mathrm{P}\to\mathrm{P}') = -qV_{\mathrm{PP}'} = -q(V_{\mathrm{P}'} - V_{\mathrm{P}}) = -q\,dV \qquad (25.35)$$

Figure 25.26 To determine the component of electric field along an axis, we calculate the electrostatic work done on a charged particle as the particle is moved over a short segment along that axis.



because the potential difference between P and P′ is $(V + dV) - V = dV$. On the other hand, we also know that the electrostatic work done on the particle is equal to the scalar product of the electric force exerted on the particle and the force displacement $d\vec{r}_F = d\vec{s}$:

$$W_q(\mathrm{P} \to \mathrm{P}') = \vec{F}_q^E \cdot d\vec{s} = (q\vec{E}) \cdot d\vec{s}$$
$$= q(\vec{E} \cdot d\vec{s}) = qE \cos\theta \, ds, \tag{25.36}$$

where we have assumed that the force displacement is small enough that $\vec{E}$ can be considered constant between P and P′. Note that θ is the angle between $\vec{E}$ and the s axis, so $E \cos\theta$ is the component of the electric field along the s axis. We can write $E \cos\theta = E_s$, and equating the two expressions for the electrostatic work done on the particle, Eqs. 25.35 and 25.36, we get

$$-q \, dV = qE_s \, ds \tag{25.37}$$

or

$$E_s = -\frac{dV}{ds}. \tag{25.38}$$

In other words, the component of the electric field along the s axis is given by the negative of the derivative of the potential with respect to s. The faster V varies (the more closely spaced the equipotentials), the greater the magnitude of the electric field.

Equation 25.38 gives only the component of the electric field along the (arbitrary) axis s. To determine all the components of the electric field, we must repeat this procedure for each of the three Cartesian coordinates:

$$E_x = -\frac{\partial V}{\partial x}, \quad E_y = -\frac{\partial V}{\partial y}, \quad E_z = -\frac{\partial V}{\partial z}. \tag{25.39}$$

Note the partial derivatives; these are necessary because the function $V(x, y, z)$ depends on all three Cartesian coordinates. When you take the partial derivative with respect to one coordinate, just remember to keep the other coordinates constant. (For example, if $V = x^2y$, then $\partial V/\partial x = 2xy$, $\partial V/\partial y = x^2$, and $\partial V/\partial z = 0$.)

Once the components of the electric field are determined, we can write the electric field in vectorial form:

$$\vec{E} = -\frac{\partial V}{\partial x}\hat{\imath} - \frac{\partial V}{\partial y}\hat{\jmath} - \frac{\partial V}{\partial z}\hat{k}. \tag{25.40}$$

Equation 25.40 then tells us how to obtain the electric field from the potential.



25.16 Apply Eq. 25.40 to the potential you obtained for the uniform field between two parallel charged plates in Example 25.5, and verify that you get the correct expression for the electric field between the plates.

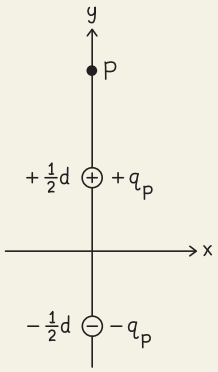
QUANTITATIVE TOOLS

EXAMPLE 25.8 The electrostatic potential and electric field due to a dipole

A permanent dipole consists of a particle carrying a charge $+q_p$ at $x = 0, y = +\frac{1}{2}d$ and another particle carrying a charge $-q_p$ at $x = 0, y = -\frac{1}{2}d$. Use the electrostatic potential at a point P on the axis of the dipole to determine the electric field at that point.

❶ **GETTING STARTED** I begin by making a sketch of the dipole (**Figure 25.27**). I place the dipole along my y axis, with the midpoint of the dipole length at the origin, so that the particles are at the coordinates given in the problem. I choose a point P on the positive y axis, and I let the y coordinate of P be y. Because the two charged particles lie on the y axis, the electric field they create at P must be directed along the y axis.

Figure 25.27



❷ **DEVISE PLAN** The potential at P is the sum of the potentials due to the individual charged particles. To calculate these potentials, I can use Eq. 25.30. Once I have calculated the potential at P, I can use Eq. 25.40 to determine the electric field.

❸ **EXECUTE PLAN** The potential at P is

$$V_P = \frac{1}{4\pi\epsilon_0}\Sigma\frac{q_n}{r_{nP}} = \frac{1}{4\pi\epsilon_0}\left[\frac{+q_p}{y-\frac{1}{2}d} + \frac{-q_p}{y+\frac{1}{2}d}\right]. \quad (1)$$

Now I use this result to determine the electric field. Because I am working with the y axis, I work with the y component of Eq. 25.40:

$$E_y = -\frac{\partial V}{\partial y} = \frac{1}{4\pi\epsilon_0}\left[\frac{q_p}{(y-\frac{1}{2}d)^2} - \frac{q_p}{(y+\frac{1}{2}d)^2}\right]. \checkmark$$

Because the electric field at P is directed along the y axis, the other components of the electric field are zero: $E_x = E_z = 0$. ✔

❹ **EVALUATE RESULT:** This is the same result as in Eq. 23.11. (Remember $k = 1/(4\pi\epsilon_0)$.)

In comparing the derivation in Example 25.8 with the derivation of the electric field in Section 23.6, note that the calculation of the electric field via the potential does not involve any vector addition. The scalar nature of the potential therefore greatly simplifies calculations.

25.17 Calculate the electric field at any point on the axis of a thin charged disk from the potential we obtained in Example 25.7. Compare your answer to the result we obtained by direct integration in Section 23.7.

Chapter Summary

Energy and work (Sections 25.1, 25.2, 25.4)

Concepts **Electric potential energy** is the potential energy associated with the relative positions of charged objects.

The **electrostatic work** done by an electric field on a charged particle as the particle moves from one position to another is independent of the path taken and depends on only the particle's initial and final positions.

The constant electric field created by stationary charged objects is sometimes called an *electrostatic field*.

Quantitative Tools If two particles carrying charges q_1 and q_2 are separated by a distance r_{12}, their **electric potential energy** U^E is

$$U^E = \frac{q_1 q_2}{4\pi\epsilon_0}\frac{1}{r_{12}}, \tag{25.8}$$

where U^E is taken as zero at infinite separation. The electric potential energy of a system of charged particles is the sum of the potential energies of all possible particle pairs. For three particles, the energy is

$$U^E = \frac{q_1 q_2}{4\pi\epsilon_0}\frac{1}{r_{12}} + \frac{q_1 q_3}{4\pi\epsilon_0}\frac{1}{r_{13}} + \frac{q_2 q_3}{4\pi\epsilon_0}\frac{1}{r_{23}}. \tag{25.14}$$

Electrostatic potential (Sections 25.2, 25.3, 25.5)

Concepts The **potential difference** between points A and B in an electrostatic field is equal to the negative of the electrostatic work per unit charge done by the field on a charged particle as it moves from A to B.

An **equipotential surface** is a surface over which the electrostatic potential is constant. The electric field lines associated with a stationary charge distribution are always perpendicular to the equipotential surfaces for the charge distribution. The electrostatic work done on a charged particle moving along an equipotential surface is zero.

An electrostatic field is directed from points of higher potential to points of lower potential. The electric force pushes positively charged particles toward regions of lower potential, whereas it pushes negatively charged particles toward regions of higher potential.

Quantitative Tools If we set the zero of potential at infinity, the potential $V(r)$ at a distance r from a particle carrying a charge q is

$$V(r) = \frac{1}{4\pi\epsilon_0}\frac{q}{r}. \tag{25.21}$$

Potential is measured in **volts:**

$$1\ \text{V} \equiv 1\ \text{J/C}. \tag{25.16}$$

If we set the zero of potential at infinity, the potential at a point P due to a group of n charged particles of charge q_n at distances $r_{n\text{P}}$ from point P is

$$V_\text{P} = \frac{1}{4\pi\epsilon_0}\sum_n \frac{q_n}{r_{n\text{P}}}. \tag{25.30}$$

The work W_q done on a charged particle q by an electric field $\vec{E}$ as the particle moves from point A to point B is

$$W_q(\text{A}\to\text{B}) = q\int_\text{A}^\text{B} \vec{E}\cdot d\vec{\ell}. \tag{25.24}$$

The potential difference V_AB between points A and B due to an electrostatic field $\vec{E}$ is

$$V_\text{AB} \equiv \frac{-W_q(\text{A}\to\text{B})}{q} = -\int_\text{A}^\text{B} \vec{E}\cdot d\vec{\ell}. \tag{25.25}$$

For any electrostatic field,

$$\oint \vec{E}\cdot d\vec{\ell} = 0. \tag{25.32}$$

Continuous charge distributions (Section 25.6)

Quantitative Tools At any point P, the potential due to an extended charged object is

$$V_P = \frac{1}{4\pi\epsilon_0}\int \frac{dq_s}{r_{sP}}, \tag{25.34}$$

where r_{sP} is the distance from the infinitesimal charge element dq_s to P and the integral is taken over the object. (See the Procedure box "Calculating the electrostatic potentials of continuous charge distributions" on page 895.)

If a thin rod of length ℓ carries a uniformly distributed charge q, the potential due to the rod at any point P on a line that runs perpendicular to the rod and passes through either end is

$$V_P = \frac{1}{4\pi\epsilon_0}\frac{q}{\ell}\ln\left(\frac{\ell + \sqrt{\ell^2 + d^2}}{d}\right),$$

where d is the distance from P to the rod.

If a thin disk of radius R carries a uniform surface charge density σ, the potential due to the disk at any point P on a line that runs perpendicular to the plane of the disk and passes through the disk center is

$$V_P = \frac{\sigma}{2\epsilon_0}(\sqrt{z^2 + R^2} - |z|),$$

where z is the distance from P to the disk.

Calculating the electrostatic field from the potential (Section 25.7)

Quantitative Tools If we know the potential due to a charge distribution, we can use the equation

$$\vec{E} = -\frac{\partial V}{\partial x}\hat{\imath} - \frac{\partial V}{\partial y}\hat{\jmath} - \frac{\partial V}{\partial z}\hat{k} \tag{25.40}$$

to calculate the value of the electrostatic field due to the distribution at any point P(x, y, z).

We can also calculate individual components of the electric field from the potential. For example, $E_x = -\partial V/\partial x$ and $E_r = -\partial V/\partial r$.

Questions and Problems

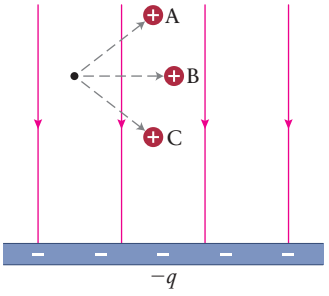
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

25.1 Electric potential energy

1. What orientation of an electric dipole in a uniform electric field has the greatest electric potential energy? What orientation has the least? (Let the system comprise both the electric dipole and the sources of the uniform electric field.) •
2. Identical positively charged objects A, B, and C are launched with the same initial speed from the same position above a negatively charged sheet that produces a uniform electric field. The directions of the initial velocities are shown in Figure P25.2. Assuming the objects do not interact with one another, rank them in order of their speeds at the instant they reach the sheet, greatest speed first. •

Figure P25.2

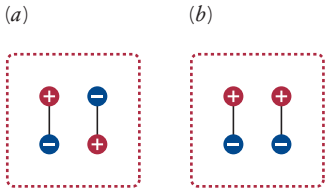


3. You release three balls simultaneously from the same height above the floor. The balls all carry the same quantity of surplus positive charge, but they have different masses: 1 kg, 2 kg, and 3 kg. In addition to the gravitational field due to Earth, there is a uniform electric field

directed downward. Assume that the balls are far enough apart that they have negligible influence on one another and that air resistance can be ignored. (*a*) Which ball has the greatest speed when it reaches the floor? (*b*) If all three were negatively charged and at least one will hit the floor, which ball would reach the floor first? ••

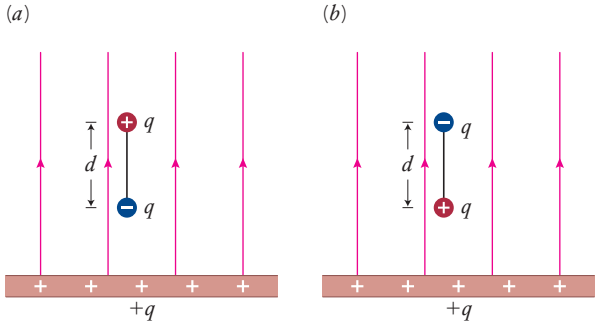
4. A proton, a deuteron (a hydrogen nucleus containing one proton and one neutron), and an alpha particle (a helium nucleus consisting of two protons and two neutrons) initially at rest are all accelerated through the same distance in the uniform electric field created by a very large charged plate. Compare their final (*a*) kinetic energies, (*b*) momentum magnitudes, (*c*) speeds, and (*d*) time intervals needed to cover the distance. (*e*) How do the changes in electric potential energy compare? ••
5. A proton, a deuteron (a hydrogen nucleus containing one proton and one neutron), and an alpha particle (a helium nucleus consisting of two protons and two neutrons) initially at rest are all accelerated for the same time interval in the uniform electric field created by a very large charged plate. Compare their final (*a*) kinetic energies, (*b*) momentum magnitudes, (*c*) speeds, and (*d*) distances traveled. (*e*) How do the changes in electric potential energy compare? ••
6. Consider an isolated system of two identical electric dipoles as in Figure P25.6. For which orientation is the electric potential energy smaller? ••

Figure P25.6



7. A dipole carrying charges $+q$ and $-q$ separated by a distance d is rotated 180° in a uniform electric field (Figure P25.7). What is the change in electric potential energy associated with this rotation? (Consider a system comprising both the dipole and the sources of the electric field.) ••

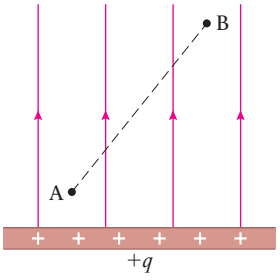
Figure P25.7



25.2 Electrostatic work

8. In an electrostatic field, path 1 between points A and B is twice as long as path 2. If the electrostatic work done on a negatively charged particle that moves from A to B along path 1 is W_1, how much work is done on this particle if it later goes from A to B along path 2? •
9. A point charge $q_1 = +5.70\ \mu\text{C}$ is fixed at the origin of a coordinate plane. How much electrostatic work is done by the electric force on a second point charge $q_2 = -3.40\ \mu\text{C}$ when it moves from the point $x = 25.0$ cm to the point $x = 37.0$ cm? •
10. A pair of electrons is fixed at certain positions (e_1 at the origin while e_2 is at point *a*) and carries an electric potential energy of 4.8×10^{-5} J. To move the second charge electron to a different location (point *b*) it requires -2.1×10^{-5} J amount of work. Find the electric potential energy of the pair of electrons when the second electron is at point *b*. •
11. Points A and B are on the same electric field line. If an electron is released and moves from A to B, (*a*) is the field directed from A to B or from B to A? (*b*) Is the potential difference between points A and B positive or negative? ••
12. In three separate experiments, an object is moved from point A to point B in the uniform electric field of a large charged plate (Figure P25.12). The object is at rest in both the initial and final positions. Object 1 carries charge $+q$ and has mass m; object 2 carries charge $+q$ and has mass $2m$; object 3 carries charge $-q$ and has mass m. Rank the objects in order of the changes their displacements produce in the electric potential energy of the object-plus-plate system, greatest change first. ••

Figure P25.12

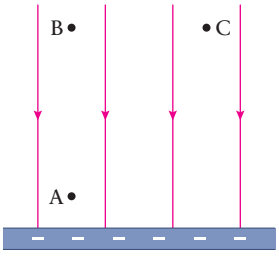


13. Electrostatic work W is done on a charged particle as the particle travels from point A to point B in an electric field. You then apply a force to move the particle back to A, increasing its kinetic energy by an amount equal to $2W$. How much work did you do? ••
14. Points A, B, and C form the vertices of a triangle in a nonuniform electrostatic field. The electrostatic work done on a particle of charge q as the particle travels from A to B is W_{AB}, and that done on the particle as it travels from A to C is $W_{AC} = -W_{AB}/2$. How much electrostatic work is done on the particle as it travels from B to C? ••
15. In the presence of an electrostatic field, you find that you must do positive work on an electron to move it from point A to point B without changing its kinetic energy. (*a*) Considering just the electron as our system, has the system's electric potential energy increased, decreased, or remained the same? (*b*) What if your system consists of

both the electron and the charged particles that are the source of the electrostatic field? (*c*) Is the electrostatic potential difference between A and B positive, negative, or zero? ●●

16. In the presence of an electrostatic field, you find that you must do positive work on a proton to move it from point A to point B without changing its kinetic energy. (*a*) Considering just the proton as your system, have you increased, decreased, or left unchanged its electric potential energy? (*b*) What if your system consists of both the proton and the charged particles that are the source of the electrostatic field? (*c*) Is the potential difference between A and B positive, negative, or zero? ●●
17. An electron moves from point A to point B under the influence of an electrostatic field in which the potential difference between A and B is negative. (*a*) If your system includes both the electron and the particles generating the electrostatic field, does the system's electric potential energy increase, decrease, or remain the same? (*b*) Does the electron's kinetic energy increase, decrease, or remain the same? (*c*) How would the answers in parts *a* and *b* change if your system included only the electron? ●●
18. A proton moves from point A to point B under the influence of an electrostatic field in which the potential difference between A and B is negative. (*a*) If your system includes both the proton and the particles generating the electrostatic field, does the system's electric potential energy increase, decrease, or remain the same? (*b*) Does the proton's kinetic energy increase, decrease, or remain the same? (*c*) How would the answers in parts *a* and *b* change if your system included only the proton? ●●
19. Four objects are moved in various ways in the electrostatic field of Figure P25.19. Rank the following motions in order of the amount of electrostatic work done on the object, smallest amount first: (*a*) an object carrying charge $+q$ is moved from A to B, (*b*) an object carrying charge $+q$ is moved from A to C, (*c*) an object carrying charge $+q$ is moved from B to C, (*d*) an object carrying charge $-q$ is moved from A to B. ●●

Figure P25.19

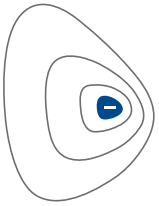


20. Points A, B, C, and D are at the corners of a square area in an electric field, with B adjacent to A and C diagonally across from A. The potential difference between A and C is the negative of that between A and B and twice that between B and D. In terms of the potential difference between A and B, what is the potential difference (*a*) between B and C, (*b*) between C and D, and (*c*) between A and D? ●●●

25.3 Equipotentials

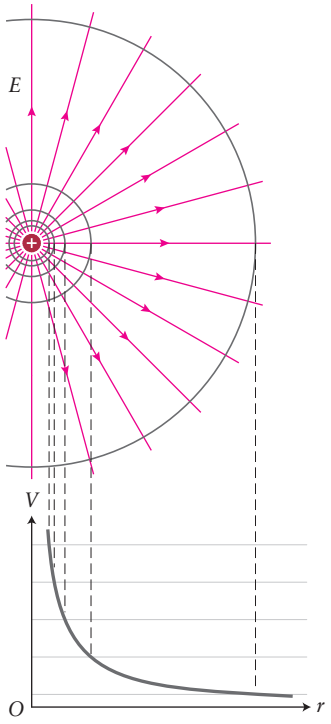
21. Can equipotential lines ever cross? ●
22. Can you draw an equipotential surface through a position where the electric field vanishes? ●
23. Describe the equipotential surfaces associated with an infinitely long charged wire that has a uniform linear charge density. ●
24. Sketch some equipotential lines for the electric field line pattern of an electric dipole. ●●
25. Sketch some electric field lines and equipotential lines associated with two identical charged particles spaced horizontally. ●●
26. Some equipotential lines surrounding a negatively charged object are shown in Figure P25.26, where the potential difference between any two adjacent lines is the same. (*a*) In which region is the electric field magnitude greatest? (*b*) What is the direction of the electric field in that region? (*c*) Which equipotential line has the greatest potential? ●●

Figure P25.26



27. You determine that it takes no (net) electrostatic work to move a charged object from point A to point B along a particular path. (*a*) Can you conclude that A and B are on the same equipotential surface? (*b*) Can you say that the path is part of an equipotential surface that includes A and B? Justify your answers. ●●
28. Figure P25.28 shows that the spacing between equipotential surfaces surrounding a charged particle increases with radial distance r away from the particle. (*a*) How would the potential have to depend on r in order to make the spacing between equipotential surfaces uniform? (*b*) How would the electric field depend on r in that case? ●●

Figure P25.28



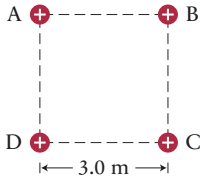
25.4 Calculating work and energy in electrostatics

29. Particle A, located at the origin, carries a charge of 6.0 nC. (*a*) What is the electrostatic potential (relative to zero potential at infinity) at a position $r = 8.0$ cm from particle A? (*b*) With particle A held in place, how much work must be done by an external force to bring particle B, also carrying a 6.0-nC charge, from infinity to a distance $r = 8.0$ cm from particle A? (*c*) Particle B is returned to infinity, and particle A is moved to $r = 8.0$ cm and held there. How much work must be done by an outside agent to bring particle B from infinity to the origin? •

30. An electron is released from rest in a uniform electric field of magnitude 4.0×10^3 N/C. What is the electrostatic work done by the electric field on the particle after it has traveled 6.0 m? ••

31. Particles A, B, C, and D in Figure P25.31 each carry a charge of magnitude 3.0 nC. Calculate the electric potential energy for the charge distribution in this 3.0-m square (*a*) if all the charges are positive; (*b*) if B, C, and D are positive and A is negative; (*c*) if B and C are positive and A and D are negative; and (*d*) if A and C are positive and B and D are negative. ••

Figure P25.31



32. An electron and a proton are fixed at the vertices of an equilateral triangle with sides of length *a*. (*a*) What is the potential (relative to zero at infinity) at the apex of the triangle. (*b*) Find the work required to move a particle with charge *Q* from infinity to the center of the line joining the fixed electron and proton? ••

33. Six particles, each carrying a charge of 2.0 nC, are equally spaced along the equator of a sphere that has a radius of 0.40 m. What is the electrostatic potential (relative to zero at infinity) (*a*) at the center of the sphere and (*b*) at either pole of the sphere? ••

34. An electron and a proton are held 2 m apart. The midpoint between the electron and the proton sets the origin of a coordinate system, with the *x* axis pointing toward the proton. (*a*) A second electron is moved to the origin from infinity. How much work does this take? (*b*) Keeping the second electron fixed at the origin, a third electron is sent toward the origin with a velocity of 400 m/s from the location (0 m, 10 m). How close does this electron get to the origin? ••

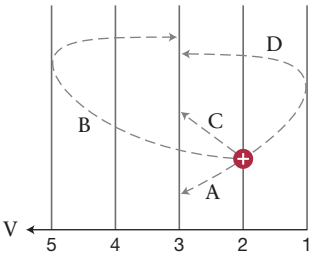
35. Four objects, each carrying a charge of magnitude *q*, are placed at the corners of a square measuring *d* on each side. Two of the objects are positively charged, and two are negatively charged, with like-charged objects placed at opposite corners of the square. (*a*) Is the electric potential energy associated with this charge distribution positive, negative, or zero? (*b*) Confirm your answer to part *a* by calculating the electric potential energy. ••

25.5 Potential difference

36. (*a*) What is the potential at a point 50 cm away from a point charge with a magnitude of 3.0 nC? (*b*) How would the potential change if you measure the potential at a point twice the distance from the charge? ••

37. Four possible paths for a positively charged object traveling from a 2-V equipotential to a 3-V equipotential are shown in Figure P25.37. Rank the paths in order of the electrostatic work done on the object along each path, greatest value first. •

Figure P25.37



38. A conducting sphere with a diameter of 8.00 cm carries a charge of 3.5 nC that is uniformly distributed. Calculate (*a*) the surface charge density of the sphere and (*b*) the electric potential of the sphere relative to the potential far away. ••

39. Two point charges $q_1 = +Q$ and $q_2 = -3Q$ are placed along the *x*-axis, 10 cm apart (q_1 at the origin, q_2 at $x = +10$ cm). (*a*) What is the electric potential at the center of the line connecting the two point charges? (*b*) Could there be a point along the *x*-axis at which the potential is zero? If yes, at what location? If no, why not? •••

40. A particle carrying charge +10.00 nC sets the origin of a rectangular coordinate system. Taking the electrostatic potential to be zero at infinity, locate the distance from the origin of the equipotential surfaces at 20 V, 40 V, 60 V, 80 V, and 100 V. Are these surfaces equally spaced? ••

41. A particle carrying charge $+q$ sets the origin of a rectangular coordinate system. A particle carrying charge $-2q$ is located on the *x* axis at $x = d$. (*a*) Taking zero potential at infinity, at what locations on the *x* axis is the electrostatic potential zero? (*b*) At what locations on the *y* axis is the potential zero? (*c*) Answer the questions of parts *a* and *b* for the case where the object at $x = d$ carries charge $+2q$. ••

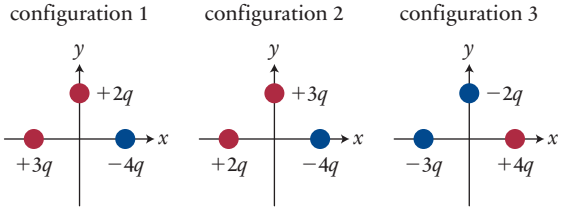
42. A positively charged infinite sheet has a surface charge density of $+5.0$ nC/m^2. What is (*a*) the electric field at a distance 0.10 m from the sheet and (*b*) the potential difference between two points located at distances of 0.30 m and 0.50 m from the sheet? (*c*) A +3.0 nC charged particle is initially at a distance 0.50 m from the sheet. How much work must an external force do to move this particle to a distance of 0.30 m from the sheet? (*d*) Taking the

potential to be zero at the sheet, at what distances from the sheet are the equipotential surfaces having potential of magnitudes 20 V, 40 V, and 60 V? ●●

43. An infinite sheet has a uniform positive surface charge. Points A and B are located at distances of 5.0 cm and 10.0 cm from the sheet, respectively. (*a*) Is the potential difference $V_B - V_A$ positive or negative? If the magnitude of this potential difference is 20.0 V, what is (*b*) the electric field in the region containing A and B and (*c*) the numerical value of the charge density on the sheet? ●●

44. Figure P25.44 shows three configurations of charged particles. All the particles are the same distance from the origin. Rank the configurations in terms of (*a*) the electrostatic potentials at the origin, greatest first, and (*b*) the electric potential energies of the three-particle system, greatest first. ●●

Figure P25.44



45. Two parallel conducting plates are 0.10 m apart and carry equal and opposite charges. They are large enough relative to the 0.10-m separation distance that we can assume the electric field between them is uniform. The potential difference between them is 500 V. An electron is released from rest at the negative plate. (*a*) Draw the electric field between the plates, indicating the sign of the charge on each plate and which plate is at the higher electrostatic potential. (*b*) What is the magnitude of the electric field between the plates? (*c*) What is the change in the potential energy of the system comprising the plates and the electron when the electron moves from the negative plate to the positive plate? (*d*) How much work is done by the electric field on the electron as it makes this trip? (*e*) What is the electron's kinetic energy when it reaches the positive plate? ●●

46. Two parallel conducting plates carry equal and opposite charges. The plates are large relative to the distance between them. The potential difference between them is 15.0 V, and the separation between them is 2 cm. (*a*) What is the numerical value of the surface charge density on the plates? (*b*) How much work is done by the electric field on an electron as it moves from the negative plate to the positive plate? ●●

47. Particle 1 carrying charge $+q_1$ is placed on the x axis at $x = -d$. Particle 2 carrying some unknown charge q_2 is placed somewhere on the x axis. The potential energy associated with the charges is $+q_1^2/(2\pi\epsilon_0 d)$, and the electrostatic potential due to the charges is $+q_1/(\pi\epsilon_0 d)$ at the origin and zero at an infinite distance. (*a*) What is the sign of the charge on particle 2? (*b*) What is the magnitude of the charge on particle 2, and where on the x axis was particle 2 placed? Determine all possible answers. ●●

48. Two charged particles are placed near the origin of an xyz coordinate system. Particle 1 carries charge $+q$ and is on the x axis at $x = +d$. Particle 2 carries an unknown charge and is at an unknown location. The magnitude of the electric field at the origin is $q/(2\pi\epsilon_0 d^2)$, and the electrostatic potential at the origin, relative to $V = 0$ at infinity, is $+3q/(4\pi\epsilon_0 d)$. (*a*) If particle 2 is on the x axis, what are its charge and location? (*b*) If particle 2 is on the y axis, what are its charge and location? Determine all possible solutions. ●●●

25.6 Electrostatic potentials of continuous charge distributions

49. A thin disk of radius $R = 5$ cm has uniform surface charge density $\sigma = 8$ nC/m^2. Calculate the potential (relative to zero at infinity) on the axis that runs through the disk center and is perpendicular to the disk face at distances (*a*) 5.0 mm, (*b*) 2 cm, and (*c*) 1 m. ●

50. A point charge $q = +5$ nC is at the center of a thick conducting spherical shell of internal radius 10 cm and external radius 12 cm. (*a*) What is the magnitude and direction of the electric field at the inner and outer surfaces of the shell? (*b*) What is the electrostatic potential on the inner and outer surfaces of the shell, and at a point located at 5 cm from the center of the shell, relative to zero at infinity? ●

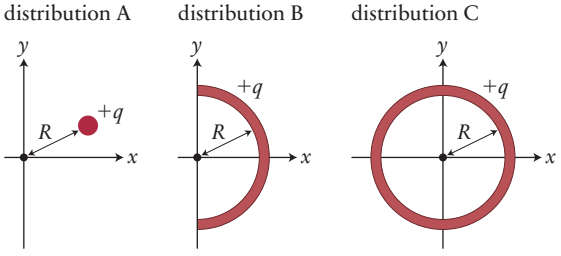
51. An infinitely long positively charged wire has a linear charge density $\lambda = 3$ nC/m. Calculate the electrostatic potential at distances of (*a*) 1.0 m, (*b*) 2.0 m, and (*c*) 6 m from the wire, assuming zero potential 4 m from the wire. ●●

52. A particle carrying charge $+q$ is placed at the center of a thick-walled conducting shell that has inner radius R and outer radius $2R$ and carries charge $-4q$. A thin-walled conducting shell of radius $3R$ carries charge $+4q$ and is concentric with the thick-walled shell. Defining $V = 0$ at infinity, calculate all distances from the particle at which the electrostatic potential is zero. ●●

53. Four uniformly charged nonconducting rods, each of length ℓ and carrying charge $+q$, are arranged into a square of side length ℓ. At the center of the square, what are (*a*) the magnitude of the electric field and (*b*) the electrostatic potential? ●●

54. Figure P25.54 shows three charge distributions. In A, a particle carrying a charge $+q$ is located a distance R from the origin. In B, $+q$ charge is spread uniformly over

Figure P25.54

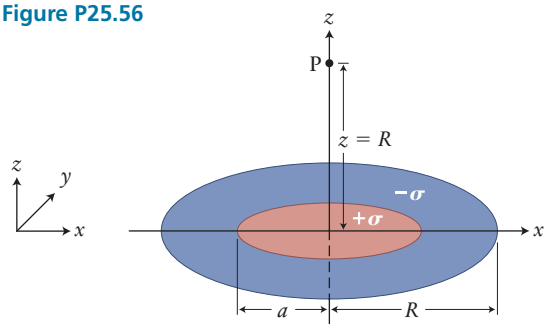


a semicircle of radius R centered at the origin. In C, $+q$ charge is spread uniformly over a circle of radius R centered at the origin. Rank these distributions in terms of (*a*) the electric field magnitude at the origin, smallest first, and (*b*) the electrostatic potential at the origin, smallest first. (*c*) Using infinity as the location where the electrostatic potential is zero, write an expression for the electrostatic potential at the origin in distribution B. ●●

55. A particle carrying charge $q_p = +10$ nC is located at $y_p = 0.030$ m on the y axis of an xy coordinate system. A nonconducting rod of length $\ell = 0.10$ m carrying charge $q_r = -10$ nC lies on the x axis, extending from $x = 0$ to $x = 0.10$ m. Write an expression for the electrostatic potential for any position along the y axis. ●●

56. A disk of radius R has positive charge uniformly distributed over an inner circular region of radius a and negative charge uniformly distributed over the outer annular (ring-shaped) region (Figure P25.56). The surface charge density on the inner region is $+\sigma$, and that on the annular region is $-\sigma$. The electrostatic potential at P, located on the central perpendicular axis a distance $z = R$ from the disk center, is zero. What is a in terms of R? ●●●

Figure P25.56



57. The surface charge density on a nonconducting disk of radius R varies with the radius as $\sigma(r) = cr$, where c is a positive constant $(c > 0)$. Derive an expression for the electrostatic potential as a function of position x along an x axis that runs through the disk center and is perpendicular to the disk face, with the disk center at $x = 0$ on this axis. ●●●

58. A very long, solid, positively charged cylinder has a radius R and is made of a nonconducting material. The nonuniform volume charge density is given by $\rho = +ar$, where r is the radial distance away from the long central axis of the cylinder. Calculate the difference in electrostatic potential between a location on the axis and a location on the outer surface of the cylinder. ●●●

25.7 Obtaining the electric field from the potential

59. The electric potential in some region of space is given by $V(x, y, z) = A\left(x^2 - xy + \dfrac{z^3}{y}\right)$, where V is in volts, x, y, z are in meters, and A is a positive constant. Determine the magnitude of the electric field in this region. ●●

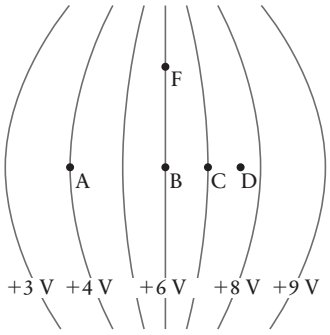
60. The electrostatic potential at a particular coordinate is given by $V(x, y) = 3x^2 - 2xy + 5y^2$. Obtain the expression for the electric field. ●

61. A particle carrying a charge of 6 nC is at the center of a thick conducting spherical shell of inner radius 9 cm and outer radius 10 cm. (*a*) What is the surface charge density on the outer surface of the shell? (*b*) What is the electrostatic potential (relative to zero at infinity) on the outer surface of the shell? (*c*) Compute the potential difference ΔV between the outer surface of the shell and a point A located at a distance $d = 1$ mm outside it. (*d*) Use the approximate expression $E = \Delta V/d$ to calculate the electric field at the surface of the shell. Compare your expression with what you would have obtained by considering the shell as an infinitely wide conducting plane carrying the same charge density. (*e*) Evaluate the error you made compared to the exact result for the electric field at A. ●●

62. Two particles, each carrying positive charge q, are on the y axis, one at $y = +a$ and the other at $y = -a$. (*a*) Calculate the potential for any point on the x axis. (*b*) Use your result in part *a* to determine the electric field at any point on the x axis. ●●

63. Figure P25.63 shows a two-dimensional slice through a set of equipotential surfaces. (*a*) At which of the locations marked A, B, C, and D is the magnitude of the electric field greatest? (*b*) Is the magnitude of the electric field greater at B or at F? (*c*) Sketch some electric field lines in relation to these equipotential surfaces. ●●

Figure P25.63



64. A particle carrying 3.00 nC of positive charge is at the origin of a rectangular coordinate system, and a particle carrying 3.00 nC of negative charge is on the x axis at $x = 6.00$ m. (*a*) Calculate the potential on the x axis at $x = 3.00$ m, relative to zero potential at infinity. (*b*) Calculate the electric field on the x axis at $x = 3.00$ m. (*c*) Calculate the potential on the x axis at $x = 3.01$ m. (*d*) Compute $\Delta V/\Delta x$ for ΔV between $x = 3.00$ m and $x = 3.01$ m, and $\Delta x = 0.01$ m, and compare your result with your result in part *b*. ●●

65. A ring of radius a carrying a uniformly distributed positive charge q is located in the yz plane of a rectangular coordinate system, with the center of the ring at the origin. (*a*) Sketch a graph of the electrostatic potential $V(x)$ as a function of distance along the x axis. (*b*) At what position is $V(x)$ a maximum? (*c*) What is E_x at this position? ●●●

66. Consider a spherically symmetrical distribution of charged particles. The magnitude of charge for the distribution is q, and the radius of the distribution is R. The electrostatic potential varies with radial distance r away from the center of the distribution. For $r > 2R$, $V(r) = -q/(4\pi\epsilon_0 r)$;

for $R < r < 2R$, $V(r) = -q/(8\pi\epsilon_0 R)$; and for $r < R$, $V(r) = -3q/(4\pi\epsilon_0 r) + 5q/(8\pi\epsilon_0 R)$. (*a*) What is the electric field in these three regions? Use $V(\infty) = 0$ in your calculations. (*b*) Describe a set of objects that could produce an electrostatic potential that behaves in this way. •••

Additional Problems

67. An electron travels from point A to point B in an electrostatic field and gains kinetic energy. (*a*) Is the electrostatic work done on the electron positive, negative, or zero? (*b*) Is the potential at A higher than, lower than, or the same as the potential at B? •

68. A proton travels from point A to point B in an electrostatic field and gains kinetic energy. (*a*) Is the electrostatic work done on the proton positive, negative, or zero? (*b*) Is the potential at A higher than, lower than, or the same as the potential at B? •

69. Calculate E_x for each potential: (*a*) $V(x) = a + bx$, $a = 4000$ V, $b = 6000$ V/m. (*b*) $V(x) = ax + b/x$, $a = 1500$ V/m, $b = 2000$ V·m. (*c*) $V(x) = ax - bx^2$, $a = 2000$ V/m, $b = 3000$ V/m^2. (*d*) $V(x) = -2000$ V (independent of x). •

70. Consider an isolated, uniformly charged spherical shell of radius R carrying positive charge q. Point A is on the shell, point B is a distance $2R$ from the center, point C is a distance $R/2$ from the center, and point D is at the center. Is the potential difference positive or negative (*a*) between A and B, (*b*) between A and C, (*c*) between A and D? ••

71. You are sketching equipotentials for a positively charged particle, having defined $V = 0$ at infinity. If you draw the 5.0-V equipotential as a circle with a radius of 50 mm, what is the radius of (*a*) the 10-V equipotential and (*b*) the 2.0-V equipotential? ••

72. A uniformly charged rod of length ℓ and charge $+q$ lies on the x axis of a Cartesian coordinate system, extending from the origin to $x = -\ell$. What is the electric potential due to this rod at a position on the positive x axis that is a distance d from the origin? ••

73. A solid sphere of radius R is concentric with a spherical shell that carries charge $+q_{\text{shell}}$ and has an inner radius of $2R$ and outer radius of $3R$. If the electrostatic potential at the common center of the sphere and shell is the same as the potential at infinity, what is the charge q_{sphere} on the solid sphere (*a*) if the sphere is conducting and (*b*) if the sphere is nonconducting and has a uniform charge distribution throughout it? •••

74. You really like your new job at the atomic physics lab. Your boss casually mentions that the electron in a helium ion (He^+) emits energy in the form of radiation as it jumps from an orbit that has a radius of 0.42 nm to an orbit that has a radius of 0.24 nm. You wonder if any more information is needed in order to figure out the energy carried off by the emitted radiation. If not, calculate the magnitude of that energy. ••• CR

75. You have always wondered exactly how strong is the interaction called the *strong nuclear interaction*, and you suspect that an element like uranium could make a good test case. You begin to wonder if the electric potential energy discussed in this chapter might provide a path to an answer. As you search for a starting point, you return to Example 25.4, which examines the energy associated with holding the hydrogen atom together. But there is only one proton in the hydrogen nucleus, whereas a uranium nucleus has 92 protons, separated from one another by about 2×10^{-15} m. Clearly something must be able to hold all those protons so close to one another. How strong is this interaction? ••• CR

76. Two charged objects 1 and 2 are held a distance r apart. Object 1 has mass m and charge $+2q$, and object 2 has mass $2m$ and charge $+q$. The objects are released from rest. Assume that the only force exerted on either object is the electric force exerted by the other object. (*a*) When the objects are in motion, what is the ratio of their kinetic energies K_2/K_1? When the objects are a distance $4r$ apart, what is the speed of (*b*) object 1 and (*c*) object 2? When the objects are very far apart, what is the speed of (*d*) object 1 and (*e*) object 2? •••



26

Charge Separation and Storage

CONCEPTS

26.1 Charge separation
26.2 Capacitors
26.3 Dielectrics
26.4 Voltaic cells and batteries

QUANTITATIVE TOOLS

26.5 Capacitance
26.6 Electric field energy and emf
26.7 Dielectric constant
26.8 Gauss's law in dielectrics

This chapter deals with generating and storing electric potential energy. To produce charged objects, positive and negative charge carriers must be pulled apart and then kept separate. Work is required to pull apart charge carriers, just as work is required to stretch a spring. In each case, this work results in energy storage in the system. We now look at what kind of changes in energy are involved in the separation of positive and negative charge carriers and how charge carriers that have been separated can be stored in simple arrangements of conductors.

26.1 Charge separation

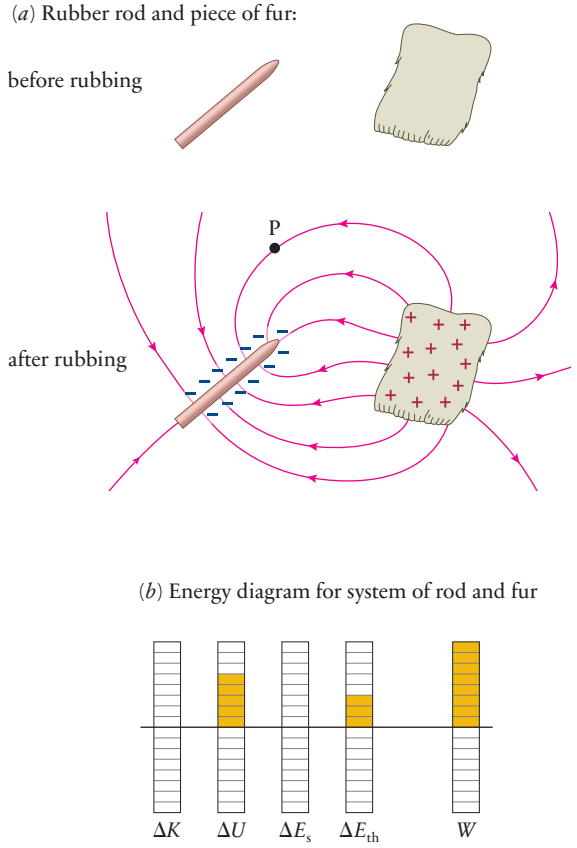
Whenever objects are "charged" (by separating strips of Scotch tape, rubbing objects against each other, using batteries, etc.), the basic phenomenon is the same: Some process (pulling, rubbing, chemical reactions) separates positive and negative charge carriers from one another. As a concrete example, consider a rubber rod and a piece of fur. If you rub the two together and then separate them, they become oppositely charged because the rod pulls electrons away from the fur: The rod ends up with a surplus of electrons and the fur with a deficit. Provided none of the electrons on the rod leak away (to the air, your hand, etc.), the magnitude of the negative charge on the rod is equal to that of the positive charge on the fur.

What is the change in energy associated with this charge separation? Consider the rubber-fur system in its initial and final states (**Figure 26.1a**). To separate the positive and negative charge carriers, they must be pulled apart against an attractive electric force, just as the ends of a stretched spring are pulled apart against an elastic force. Because work must be done on the rod-fur system, the electric potential energy of the system is greater in the final state. This energy is supplied by you while you rub the two objects together and then increase their separation. Not all of the energy you put into the system goes into electric potential energy; the friction involved in the rubbing not only produces charge separation but also heats up the rod and the fur, so part of the work you do on the system increases the thermal energy. An energy diagram for the rod-fur system is shown in Figure 26.1*b*.

26.1 Suppose you repeat the charging (starting again with uncharged rod and fur), but this time you rub longer and twice as much charge accumulates at each point on the two objects. How do the following quantities compare to what they were after the first charging: (*i*) the direction and magnitude of the electric field at point P in Figure 26.1*a*; (*ii*) the potential difference between two fixed points on the rod and the fur; and (*iii*) the electric potential energy in the rod-fur system?

Checkpoint 26.1 highlights the essence of this chapter. Be sure not to confuse *potential difference* and *electric potential energy*:

Figure 26.1 When we charge a rubber rod and a piece of fur by rubbing them together, we do work to separate charge and hence increase the electric potential energy of the system comprising the rod and fur.



- The system's electric potential energy depends on the configuration of the positive and negative charge carriers in the system.
- The potential difference between points on the rod and the fur is a measure of the electrostatic work done on a particle carrying a unit of charge (not part of the system) while moving between those points.

In the next section we examine the proportionality between potential difference and charge separation in detail. In this section we concentrate on the relationship between charge separation and electric potential energy.

The crucial point to take away from Checkpoint 26.1 is:

Positive work must be done on a system to cause a charge separation of the positive and negative charge carriers in the system. This work increases the system's electric potential energy.

26.2 Will doubling the separation distance between the charged rod and fur in Figure 26.1*a* increase, decrease, or not change the potential energy of the rod-fur system?

The amount of stored electric potential energy depends on the amount of charge that is separated and the distance that separates the charge carriers. More charge or a greater separation means more electric potential energy is

CONCEPTS

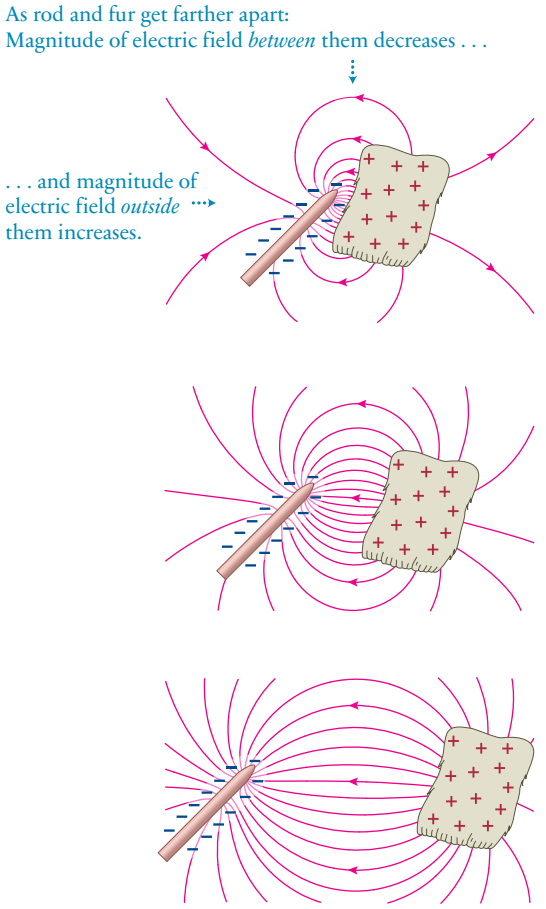
stored. These arguments apply to *all* devices that separate charge, such as Van de Graaff generators (see the box below) and batteries (see Section 26.4). Once electric potential energy has been generated by separating charge carriers, this energy can be used for other processes, such as lighting a lamp, operating a radio, and so on.

Every **charge-separating device** (or **charging device**) has some mechanism that moves charge carriers *against* an electric field—a process that requires work to be done on the system of charge carriers. For the rod-fur system, this work is mechanical and is supplied by the person doing the rubbing and separating the objects. In a battery, chemical reactions drive charge carriers through a region where the electric field opposes their motion.

26.3 If you include the person doing the rubbing in the system considered in Figure 26.1, what is the resulting energy diagram?

Where is the electric potential energy of the rod-fur system stored? As you may recall from Section 7.2, potential energy is stored in reversible changes in the configuration of interacting components of a system. Electric potential energy, therefore, is associated with the configuration of the charge carriers in a system. A look at the field line pattern suggests an alternative view, however. **Figure 26.2** shows how the electric field line pattern changes as the distance between the charged rod and the fur increases. Note how more of the space around the system becomes filled with field lines

Figure 26.2 Change in the electric field line pattern as the distance between the rod and fur increases.



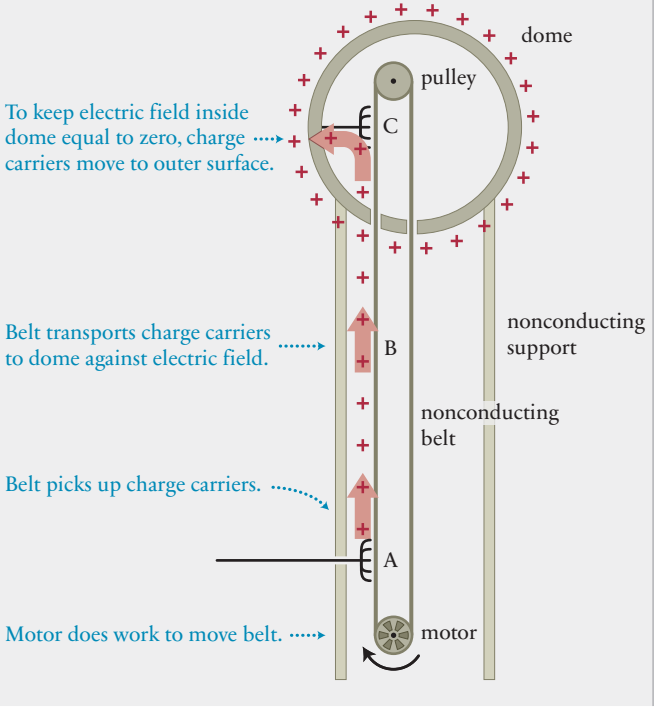
Van de Graaff generator

Figure 26.3 shows a schematic diagram of a Van de Graaff generator—a mechanical device invented in the 1930s by Robert J. Van de Graaff to separate large amounts of electrical charge. The basic principle is extremely simple: A nonconducting belt delivers charge carriers to a hollow conducting dome that rests on a nonconducting support. Machines of this type are used to generate the very large potential differences required in particle accelerators and for the generation of x rays.

Operation of the generator involves three important steps. The first step is a transfer of charge carriers to the belt at A. This transfer can be done by literally "spraying" charged particles onto the belt or simply by rubbing the rubber belt against some appropriate material.

The second step transports the charge carriers to the dome. This step is possible because the belt is nonconducting, so the charge carriers are not mobile—they are stuck to the belt, which is driven by a motor around a pulley inside the dome. The motor must do work on the charge carriers to move them against the electric field of the dome. (In the example shown in the diagram, positive charge carriers at B

Figure 26.3 Schematic diagram of a Van de Graaff generator.



CONCEPTS

must be transported upward against the downward electric field of the positively charged dome.)

The third step transfers the charge carriers from the belt onto the dome. As we saw in Section 24.5, the electric field inside a hollow conductor is always zero and any charge inside a conductor moves toward the outer surface. Therefore, once the charge carriers are inside the dome, they tend to move to the outer surface of the dome. For this purpose, a comb of conducting needles is placed close to the belt at C. If the charge carriers on the belt are electrons, the electrons hop onto the comb and move via the connecting wire toward the outside of the metal dome, causing the dome to acquire a negative charge. Alternatively, the charge carriers on the belt can be positively ionized air molecules, in which case electrons in the comb are attracted toward the ions. These electrons then jump from the comb onto the belt, neutralizing the ions on the belt while leaving a positive charge behind on the outside of the dome.



This huge double Van de Graaff generator, located at the Boston Museum of Science, was built in South Dartmouth, Massachusetts, for the MIT Physics Department in 1931. The oppositely charged pair of generators can produce potential differences of up to 5,000,000 V between the two 4.5-m domes.

(indicating that the magnitude of the electric field there increases), while the density of the field lines between the rod and the fur decreases. Thus, as the distance between the rod and the fur changes, the electric potential energy changes and the electric field changes.

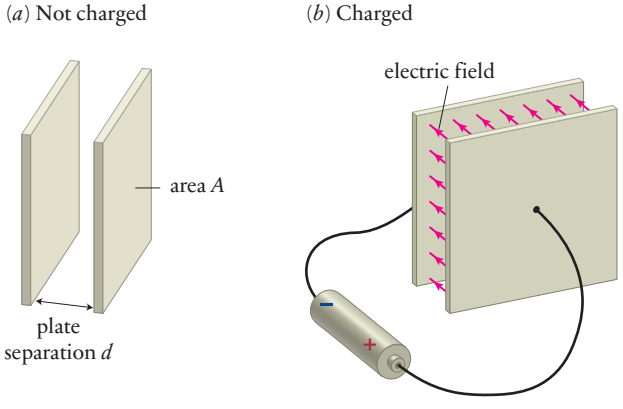
As we shall see later in this chapter, we can relate a change in the electric potential energy of a system to a change in the system's electric field (integrated over all of space), which suggests that the electric potential energy of a system is stored in its electric field. In other words, the electric potential energy of the rod-fur system is spread throughout the space around it. As long as the two objects are held stationary, the field is stationary, so the exact "location" of the energy is not very important because there is no way we can determine it experimentally. If we shake the charged rod or fur, however, the shaking causes a wavelike disturbance in the electric field. This disturbance propagates away from the rod or fur and carries with it energy that we can detect. For now we don't need to concern ourselves with such waves. It suffices to know that electric fields store electric potential energy.

26.2 Capacitors

Any system of two charged objects, such as the rod-fur system in the preceding section, stores electric potential energy. To study how much electric potential energy can be stored in a system of two objects, let's begin by considering the simple arrangement of two parallel conducting plates shown in **Figure 26.4*a***. A system for storing electric potential energy that consists of two conductors is called a **capacitor**; the arrangement in Figure 26.4*a* is called a *parallel-plate capacitor*.

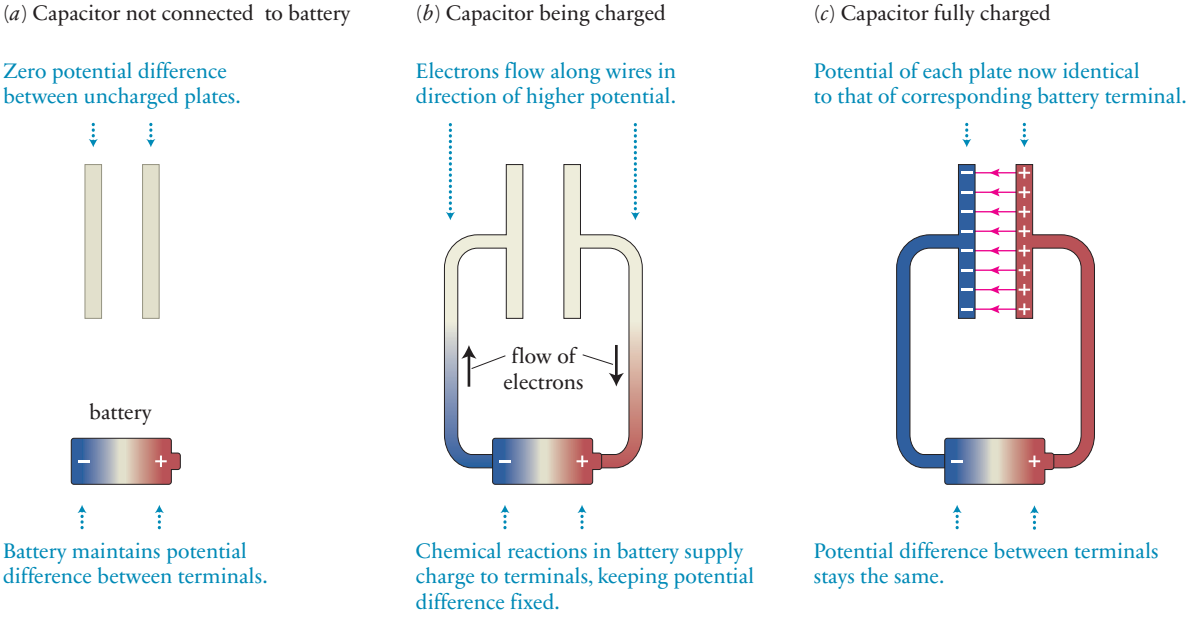
Figure 26.4*b* illustrates a simple method for charging such a capacitor. Each plate is connected by a wire to a terminal of a battery, which maintains a fixed potential difference between its terminals. **Figure 26.5** on the next page shows what happens when the connection is made between the battery and the capacitor. If the capacitor plates are far enough away from the battery, the potential difference between the plates initially is zero (Figure 26.5*a*). Immediately after the wires are connected, there is a

Figure 26.4 A parallel-plate capacitor.



CONCEPTS

Figure 26.5 Charging a capacitor. (Think of the plates as an edge-on version of those shown in Figure 26.4.)



potential difference between the ends of each wire. This difference in potential causes electrons (which are mobile in metal) in the wires to flow as indicated by the arrows in Figure 26.5*b*. A positive charge builds up on the plate connected to the positive terminal, and a negative charge of equal magnitude builds up on the other plate. As electrons leave one plate and accumulate on the other, the potential difference between the plates changes. This process continues until the potential is the same at both ends of each wire—that is, when the potential difference between the plates is equal to that between the terminals of the battery (Figure 26.5*c*). Because there is no longer any potential difference from one end of the wire to the other, the flow of electrons stops and the capacitor is said to be *fully charged*. In the process of achieving this state, the battery has done work on the electrons; this work has now become electric potential energy stored in the capacitor.

The time interval it takes to fully charge a capacitor depends on the properties of the capacitor, the battery, and the way the capacitor is connected to the battery. Typically, only a fraction of a second is needed for charging, although the time interval it takes to charge very large capacitors can be minutes (more on this in Chapter 32).

CONCEPTS

26.4 (*a*) Suppose that we disconnect the wires from the plates after the capacitor is charged as shown in Figure 26.5*c*. How does the potential difference between the plates after the wires are disconnected compare to that just before they are disconnected? (*b*) If we replace the battery in Figure 26.5 by a battery that maintains a greater potential difference between its terminals, is the magnitude of the charge on the plates greater than, smaller than, or the same as when the first battery is connected?

When a capacitor is not connected to a conductor, as in Checkpoint 26.4*a*, it is said to be *isolated*. For an isolated capacitor, the *quantity of charge on each plate* is fixed because the charge carriers have nowhere to go. In contrast, for a capacitor that is connected to a battery, the *potential difference across the capacitor* is fixed—the charge carriers on the plates always adjust themselves in such a way as to ensure that the potential difference across the capacitor is equal to that across the battery terminals.

Figure 26.6 shows the electric fields of two isolated charged parallel-plate capacitors. The field is nearly

Figure 26.6 Effect of plate separation distance in relation to plate area on the field of a parallel-plate capacitor.

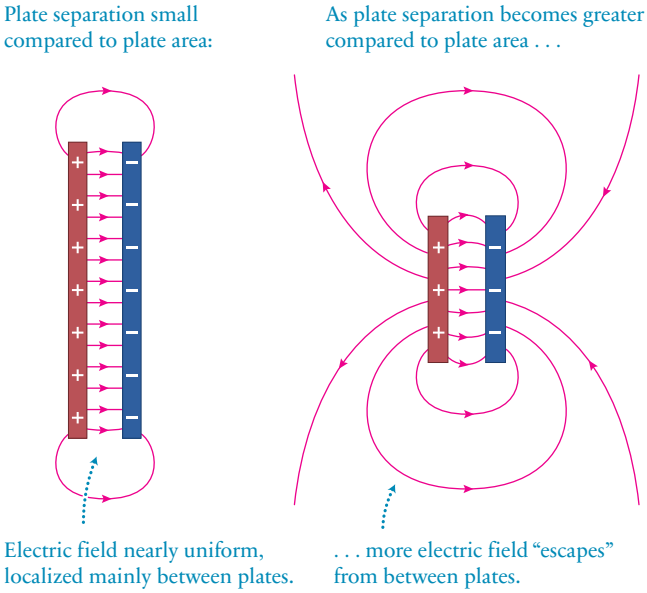


Figure 26.7 Doubling the charge on a parallel-plate capacitor.

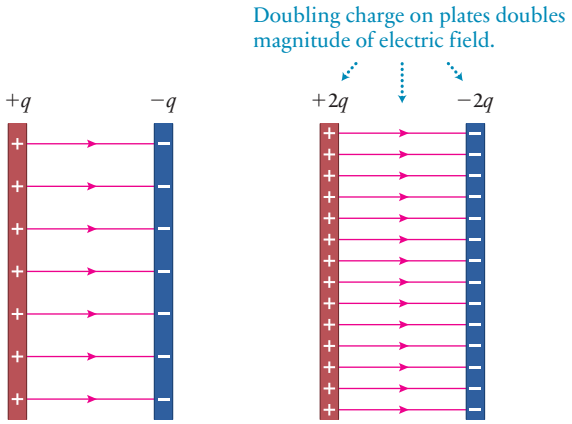
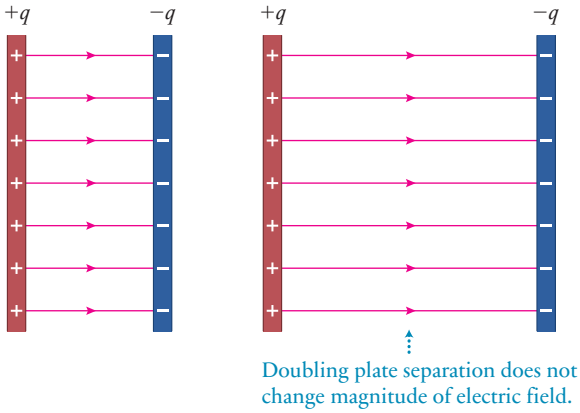


Figure 26.8 Doubling the plate separation distance of a parallel-plate capacitor.



uniform in the region between the plates, but it is nonuniform at the edges. When the spacing between the plates is small compared to the area of the plates, the effect of the nonuniform field is negligible: The electric field is confined almost entirely to the region between the plates, and for all practical purposes we can consider this field to be uniform. When discussing parallel-plate capacitors, we shall ignore the nonuniform fields at the edges and assume the electric field is entirely uniform between the plates. This simplification is justified by the geometry of most capacitors used in electronic applications.

Let us now examine the relationship between the magnitude of the charge on the plates of a parallel-plate capacitor and the potential difference between them. **Figure 26.7** shows an isolated parallel-plate capacitor carrying a positive charge $+q$ on one plate and a negative charge $-q$ on the other. If we double the magnitude of the charge on each plate, then the electric field between the plates doubles, too (see Checkpoint 23.14). Consequently, the electric force exerted on a charged particle between the plates doubles, so the electrostatic work done in moving a charged particle from one plate to the other also doubles. According to Eq. 25.15, the potential difference between the plates doubles as well. In other words, the potential difference between the plates is proportional to the magnitude of the charge on the plates.

What happens if we increase the plate separation of an isolated parallel-plate capacitor, as illustrated in **Figure 26.8**? The electric field remains the same because it is determined by the surface charge density on the plates, which doesn't change (see Checkpoint 23.14). Because the distance between the plates increases, however, the electrostatic work done in moving a charged particle from one plate to the other increases—more work is required to move the particle over a greater distance—so the potential difference between the plates increases too, confirming the result we obtained in Example 25.5.

26.5 Suppose the two capacitors in Figure 26.8 are each connected to a 9-V battery. (*a*) Which of the two capacitors stores the greater amount of charge? (*b*) If, instead of the separation distance increasing, the area of the plates of the capacitor is halved and then the capacitor is connected to a 9-V battery, does the capacitor store more charge, less charge, or the same amount of charge as before the area of the plates was halved?

As Checkpoint 26.5 illustrates, the geometry of the capacitor determines its capacity to store charge. In general:

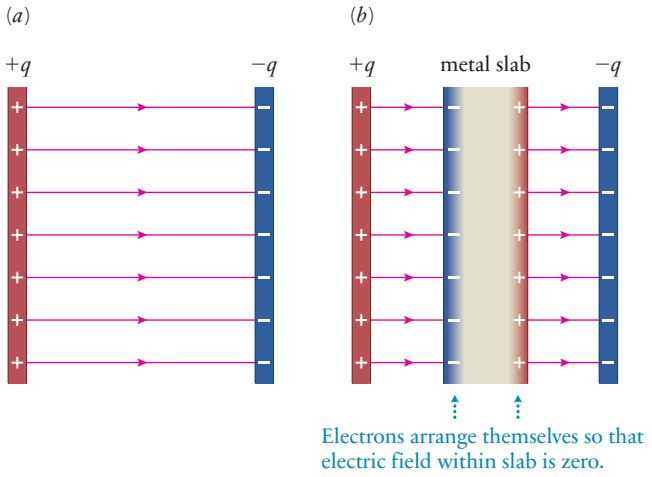
For a given potential difference between the plates of a parallel-plate capacitor, the amount of charge stored on its plates increases with increasing plate area and decreases with increasing plate separation.

Does this mean that we can increase the amount of charge stored on a parallel-plate capacitor indefinitely simply by making the plate separation infinitesimally small? The answer is *no*, because if the plate spacing is decreased while the potential difference between the capacitor plates is fixed, the charge on each plate increases and thus the magnitude of the electric field in the capacitor increases. When the electric field is about 3×10^6 V/m, the air molecules between the plates become *ionized* and the air becomes conducting, allowing a direct transfer of charge carriers between the plates. Once such a so-called **electrical breakdown** occurs, the capacitor loses all its stored energy in the form of a spark.

The electrical breakdown of air will occur between the charged dome of a Van de Graaff generator and a nearby metal object, as shown in the box on page 910. The breakdown limits the maximum potential difference across a capacitor and thus the maximum amount of charge that can be stored on the capacitor. The electric field at which electrical breakdown occurs is called the *breakdown threshold*.

CONCEPTS

Figure 26.9 Inserting a conducting slab between the plates of a parallel-plate capacitor.



The breakdown threshold can be raised by inserting a nonconducting material between the capacitor plates. As we shall see in the next section, such a nonconducting material also greatly increases the amount of charge that can be stored by a capacitor. To understand why this is so, we begin by considering a simpler situation: the insertion of a conductor between the plates of a parallel-plate capacitor.

Figure 26.9*a* shows an isolated charged capacitor. Suppose we now insert a conducting slab between the plates of this capacitor (Figure 26.9*b*). As we saw in Section 24.5, the charge carriers in the conductor rearrange themselves in such a fashion as to eliminate the field inside the bulk of the conductor.

CONCEPTS

26.6 Suppose the capacitor in Figure 26.9*a* is charged and then disconnected from the battery. (*a*) As the conducting slab is inserted in the capacitor, as in Figure 26.9*b*, does the amount of charge on the capacitor plates increase, decrease, or stay the same? (*b*) How much charge accumulates on each side of the slab, once it is inserted? (*c*) What is the potential difference across the conducting slab? (*d*) As the slab is inserted, does the magnitude of the potential difference between the capacitor plates increase, decrease, or stay the same?

As Checkpoint 26.6 demonstrates, the effect of the slab is to make the electric field in part of the space between the capacitor plates zero, thus reducing the magnitude of the potential difference between the plates for a given amount of charge on them. As the next example shows, the converse of this fact is that, for a given potential difference between its plates, the capacitor can store more energy with the slab inserted than it can without the slab. In other words, the slab increases the capacitor's capacity to store charge.

EXAMPLE 26.1 Metal-slab capacitor

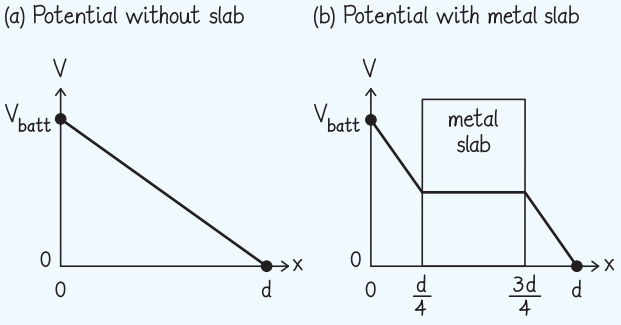
Suppose the capacitor in Figure 26.9 has a plate separation distance d and the plates carry charges $+q$ and $-q$ when the capacitor is connected to a battery that maintains a potential difference V_{batt} between its terminals. If a metal slab of thickness $d/2$ is inserted midway between the plates while the battery remains connected, what happens to (*a*) the magnitude of the electric field between the plates and (*b*) the quantity of charge on the plates?

1 GETTING STARTED I am given the plate separation distance and plate charges for a capacitor connected to a battery, and I must determine how the electric field magnitude between the plates and the quantity of charge on each plate change when a metal slab is inserted. Because the capacitor remains connected to the battery, the potential difference across the capacitor is fixed. Because the slab is conducting, the electric field inside the slab is always zero.

2 DEVISE PLAN The electric field magnitude determines the electrostatic work, which in turn determines the potential difference between the plates (which I know). To determine how the electric field magnitude changes when the metal slab is inserted, therefore, I must first determine how the potential between the plates changes when the slab is inserted. Once I know the electric field, I can determine how the slab affects the charge on the plates because I know that the magnitude of the electric field between the plates is proportional to the surface charge density σ (see Checkpoint 23.14, $E = 4k\pi\sigma$).

3 EXECUTE PLAN To compare the potentials before and after the slab is inserted, I must plot V as a function of position between the plates (**Figure 26.10**). I choose my x axis to be parallel to the electric field, with the positive plate at $x = 0$ and the negative plate at $x = d$ as the zero of potential. In the absence of the metal slab, the field is uniform between the plates, and so the potential decreases linearly from $x = 0$ to $x = d$ (Figure 26.10*a*). Because the electric field inside the slab is zero, the potential does not vary across the slab (it is an equipotential; see Section 25.3). Because the battery keeps the potential at each plate constant, the potential-versus-distance curve must take on the zigzag form shown in Figure 26.10*b*.

Figure 26.10



(*a*) Insertion of the metal slab affects the electric field between the plates because the slab is an electrical conductor and so the electric field inside it must always be zero. However, because V_{batt} is constant, I know that the electrostatic work, $W_q = F_d^E = qV_{batt}$, done on a particle carrying a charge q to move the particle from one plate to the other must be the same whether or not the slab is in place there. Because no electrostatic work is done to move the particle through the slab where the electric field is zero, the field outside the slab must be greater to make up for the smaller distance over which the particle is moved. More precisely, because the distance over which the electric field is nonzero is reduced to $d/2$, the magnitude of the field must be twice what it was before the slab was inserted. ✔

(*b*) If the field doubles, then the charge per unit area must also double. Given that the area of the plates does not change, this means that the charge on the plates must double. ✔

4 **EVALUATE RESULT** Inserting the metal slab with the battery connected is equivalent to halving the separation distance between the plates while keeping the potential difference constant. I know from Checkpoint 26.5 that, for a constant potential difference, the quantity of charge stored on a capacitor increases with decreasing plate separation, as I found.

26.7 (*a*) Does the position of the slab in Figure 26.9 affect the potential difference across the capacitor? Consider, in particular, the case in which the slab is moved all the way to one side and makes electrical contact with one of the plates. (*b*) Sketch the potential $V(x)$ as a function of x, with the slab off-center.

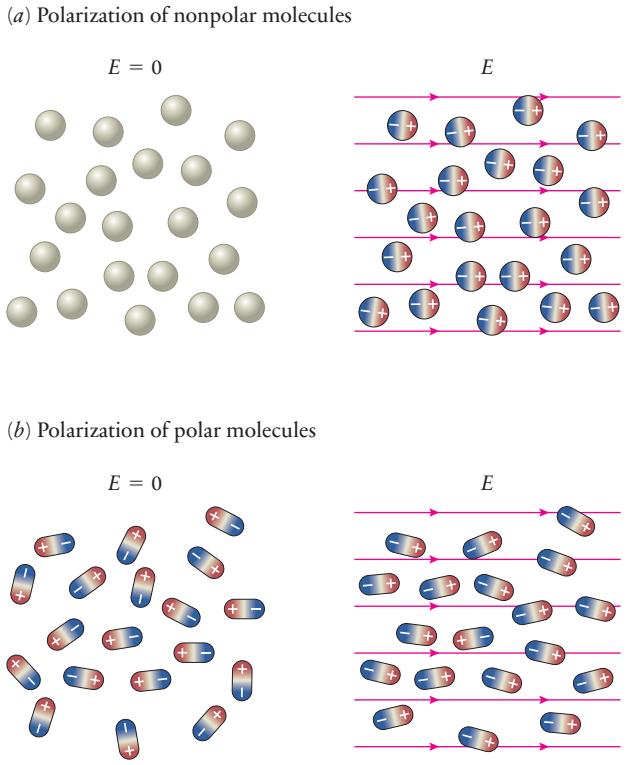
26.3 Dielectrics

As we just saw, decreasing the space inside an isolated charged capacitor where the electric field is nonzero increases its capacity to store electrical charge for a given potential difference across its plates. With a conducting slab inserted, however, the gap between either plate and the slab face nearest it is smaller than the plate-to-plate gap before the slab was inserted. Because a decreased gap with the potential difference held constant means E increases, we still have the problem of electrical breakdown.

Suppose, however, that we insert a nonconducting material—a **dielectric**—between the plates of a capacitor. As we discussed in Section 22.4, the electric field between the plates of the capacitor polarizes the dielectric. What effect does this polarization have? To answer this question we must first look in more detail at what happens in a polarized dielectric material.

We should distinguish between two general types of dielectric materials. A *polar* dielectric consists of molecules that have a permanent electric dipole moment;

Figure 26.11 Polarization of nonpolar and polar molecules in an electric field.



each molecule is electrically neutral, but the centers of its positive and negative charge distributions do not coincide (see Section 23.4). The atoms or molecules in a *nonpolar* dielectric have no dipole moment in the absence of an electric field.

Figure 26.11*a* shows the polarization of a nonpolar dielectric. In the presence of an electric field, the electrons in a nonpolar dielectric are displaced in the direction opposite to $\vec{E}$, inducing a dipole moment on each molecule.

26.8 Why are the electrons displaced in a direction *opposite* the electric field?

The polarization of a polar dielectric is shown in Figure 26.11*b*. In the absence of an electric field, the individual molecules' dipole moments are randomly aligned, so the material as a whole is not polarized. In the presence of an electric field, however, the molecular dipoles are subject to a torque (see Section 23.8) that tends to align the dipoles with the electric field, giving rise to a macroscopic polarization. In general, the polarization of polar dielectrics is much greater than that of nonpolar ones because the permanent dipole moments of the molecules in a polar dielectric are much greater than the induced dipole moments in a nonpolar dielectric.

CONCEPTS

Figure 26.12 The reason a polarized dielectric exhibits a macroscopic polarization.

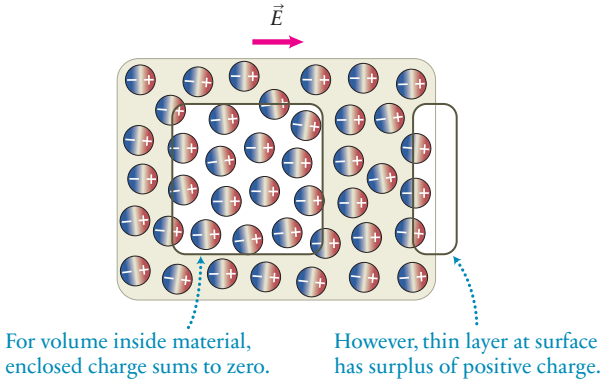
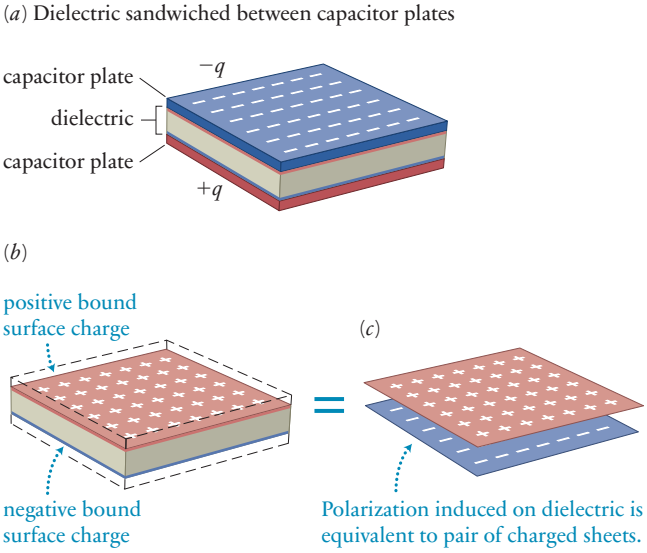


Figure 26.12 illustrates the effect of the uniform polarization of the atoms or molecules in a polar or nonpolar dielectric. The charge enclosed by any volume that lies entirely inside the polarized dielectric is zero: The positive and negative charge carriers may not coincide exactly, but on average they occur in equal numbers. However, this is not true for a small volume at the surface of the dielectric. In the volume on the right in Figure 26.12, for example, a surplus of positive charge appears at the surface of the material. Thus, a polarized dielectric has a very thin sliver of surplus positive charge on one side and a sliver containing an equal amount of surplus negative charge on the other side (see also Figure 22.21). The surface charge on either side of the polarized dielectric is said to be **bound** because the charge carriers that cause it are not free to roam around in the material. In contrast, the charge on the capacitor plates is **free** because the charge carriers that cause it can move around freely. From a macroscopic point of view, a uniformly polarized dielectric differs from an unpolarized dielectric only by the presence of this bound surface charge.

What is the effect of the bound surface charge on the electric field inside the capacitor? Consider a dielectric slab inside an isolated charged capacitor (**Figure 26.13a**). The dielectric is polarized by the electric field of the capacitor; that is, a positive bound surface charge appears on the surface near the negatively charged capacitor plate, and a negative bound surface charge appears on the other side (Figure 26.13*b*). Imagine now that we could "lock in" the polarization and consider just the slab by itself. Except for two sheets of charge at the top and bottom, the bulk of the dielectric is neutral. Thus, for all practical purposes, the polarized dielectric is equivalent to two very thin sheets carrying opposite charge (Figure 26.13*c*).

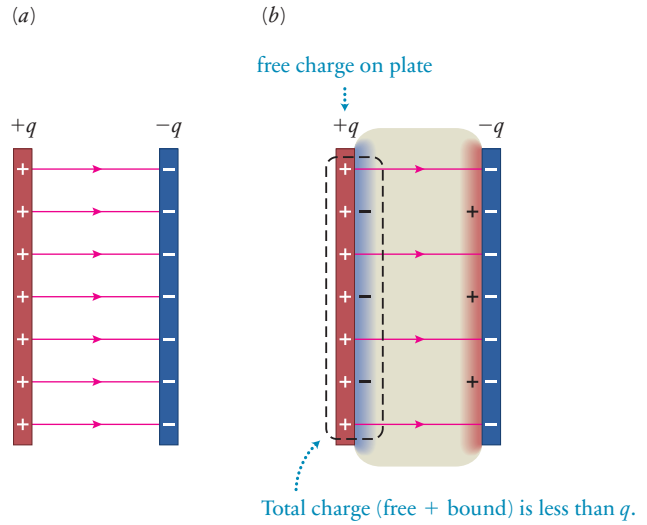
26.9 (*a*) In which direction does the electric field due to the bound surface charge point at a location above the top surface in Figure 26.13*c*? (*b*) In which direction does it point at a location between the top and bottom surfaces?

Figure 26.13 The polarization induced on a dielectric in a parallel-plate capacitor is equivalent to two thin sheets carrying opposite charge.



We can now obtain the electric field of the capacitor with the dielectric by superposition: It is equal to the electric field of the capacitor without the dielectric, plus the electric field of the polarized dielectric by itself. As you found in Checkpoint 26.9, the direction of the electric field due to the polarized dielectric is opposite that of the capacitor, so the presence of the dielectric decreases the electric field strength in the capacitor. Alternatively, we can say that each of the bound surface charges compensates for part of the free charge on the adjoining capacitor plate, so, in effect, the total charge (free and bound) on each side of the capacitor is reduced (**Figure 26.14**). This

Figure 26.14 The presence of a polarized dielectric reduces the strength of the electric field between the plates of an isolated capacitor.



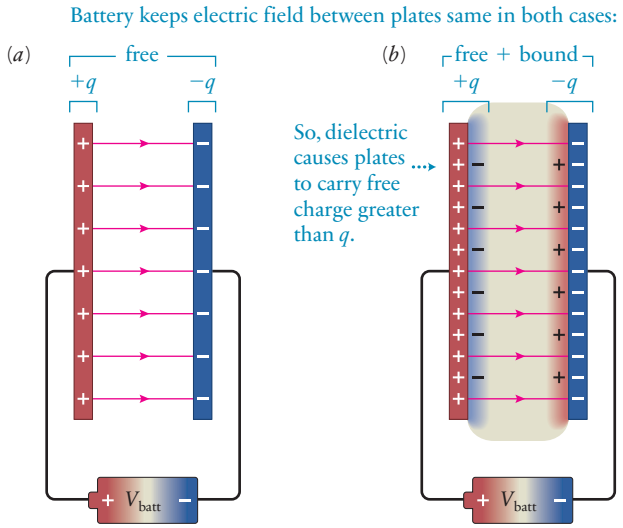
CONCEPTS

reduction in charge, in turn, gives rise to a smaller electric field inside the capacitor. For some dielectric materials, the field inside can be reduced by a factor of several thousand.

26.10 (*a*) If the magnitude of the bound surface charge on the dielectric slab in Figure 26.14*b* were equal to the magnitude of the free charge on the capacitor plates, what would be the electric field inside the capacitor? (*b*) Could the magnitude of the bound surface charge ever be *greater* than the magnitude of the free charge on the plates?

Figure 26.15 shows what happens when a dielectric-filled capacitor is connected to a battery. The battery keeps the potential difference between the capacitor plates the same regardless of the presence of the dielectric. Because the electric field in the capacitor is equal to the potential difference divided by the distance between the plates (see Example 25.5), it follows that, as long as the capacitor is connected to the battery, the electric field must be the same regardless of the dielectric. The electric field can be the same only if the distribution of charge causing the electric field is the same. In other words, regardless of the presence of the dielectric, we must have the same amount of total charge (free and bound) on each side of the capacitor. Let the magnitude of the free charge on the capacitor plates without the dielectric be q. As shown in Figure 26.15*b*, the polarization of the dielectric causes a negative bound surface charge next to the positive capacitor plate; the sum of the free charge on the conductor and the adjoining bound surface charge must still be equal to $+q$. Similarly, the sum of the negative free charge on the opposite plate and the positive bound surface charge on the adjoining dielectric still must be $-q$. Consequently, the magnitude of the free charge on each plate by itself can be much greater than without the dielectric. This extra charge is supplied by the battery.

Figure 26.15 The presence of a polarized dielectric increases the charge on the plates of a capacitor connected to a battery.



26.11 Given that the electric field is the same in both capacitors in Figure 26.15, which stores the greater amount of electric potential energy?

If the answer to Checkpoint 26.11 surprises you—after all, the electric fields are the same in the two capacitors—remember that the amount of *charge separation* is not the same. The charge on the capacitor plates polarizes the dielectric, and this polarization is the result of charge separation in the molecules of the dielectric. Thus, instead of empty space without charge separation between the capacitor plates, we now have (in addition to a greater charge on the plates) a lot of additional charge separation on the microscopic scale. Most of the energy stored in the capacitor is not due to the separation of charge on the plates, but to the separation of charge in the dielectric between the plates. The pulling apart of the positive and negative charge distributions in the dielectric increases the electric potential energy stored in the dielectric, much like stretching a spring by pulling its ends apart stores elastic potential energy. This tells us that an electric field of a given magnitude in a dielectric stores more energy than an equal field in vacuum.

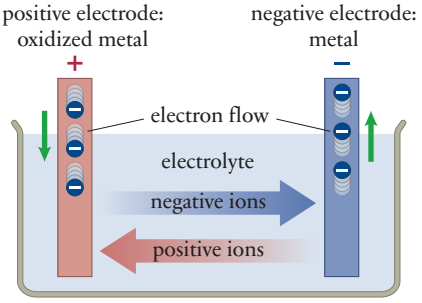
26.4 Voltaic cells and batteries

CONCEPTS

Electric potential energy is generated by separating charged particles. Earlier in this chapter we discussed two means of accomplishing such charge separation: charging by rubbing and the Van de Graaff generator. Another common way to generate electric potential energy is by means of a *voltaic cell*, the first of which was constructed by Alessandro Volta in around 1800. Assemblies of voltaic cells are called *batteries*. A standard 9-V alkaline battery, for example, consists of six 1.5-V cells connected together. While there are many types of voltaic cells and batteries, all have a common operating principle: Chemical reactions turn chemical energy into electric potential energy by accumulating electrons on one side of the cell (the negative terminal) and removing electrons from the other side (the positive terminal).

The general principle of a voltaic cell is illustrated in **Figure 26.16** on the next page. Two conducting terminals, or *electrodes*, are submerged in an *electrolyte*—a solvent that contains mobile ions. One electrode is usually

Figure 26.16 General operating principle of a voltaic cell. Electrons flow when the cell is connected to an electronic device.



made from an oxidized metal; the oxidized metal reacts by accepting positive ions from the electrolyte and electrons from the electrode. The other electrode is generally metallic; it oxidizes by taking in negative ions and giving up electrons. Because of these reactions, a surplus of electrons builds up on the metallic terminal and a deficit of electrons builds up on the oxidized-metal terminal, causing a potential difference between the two. The reactions stop when the potential difference between the electrodes reaches a certain value called the *cell potential difference*. This value is determined by the type of chemicals in the cell and is typically on the order of a few volts. As we saw in Section 25.5, this potential difference can be used to do electrostatic work on charge carriers when a battery is connected to some device—such as a light bulb, a motor, or a capacitor.

As long as the cell is not connected to anything and its chemicals do not deteriorate, the cell remains in the same state indefinitely. When the cell is connected to a capacitor or to some other device, however, the surplus of electrons is removed from the negative electrode and electrons are supplied to the positive electrode. Then the chemical reactions resume in order to maintain the cell potential difference between the terminals. As the reactions proceed, the electrolyte becomes more dilute and the compositions of the electrodes change. The cell is exhausted when all the ions in the electrolyte have been depleted. For some types of cells, the chemical reactions can be reversed by supplying a potential difference to the terminals of the cell; electric potential energy is then converted back to chemical energy. Such cells are used in rechargeable batteries, which can be reused repeatedly to store and recover electric potential energy.

As we noted in Section 26.1, any charge-separating device involves the motion of charge carriers in the opposite direction of an electric field. This is where work is done on the charge carriers and where some form of energy is converted to electric potential energy. Inside a voltaic cell, electrons must be pulled away from the positively charged terminal and deposited onto the negatively charged terminal. This process occurs at the surface of the electrodes where the chemical reactions take place. The chemical reactions move charge carriers in the opposite direction of electric forces. The chemical energy released in the reactions provides the energy necessary to move the charge carriers against the electric field. The work done per unit charge is called the **emf** (pronounced e-m-f)* of the device:

> **The emf of a charge-separating device is the work per unit charge done by nonelectrostatic interactions in separating positive and negative charge carriers inside the device.**

All charge-separating devices—batteries, voltaic cells, generators, solar cells—have some *nonelectrostatic* means to separate charge carriers and thereby create a potential difference across the terminals of the device.

26.12 As electrons leave one terminal and are added to the other, ions in the electrolyte must flow in the direction indicated in Figure 26.16 to maintain an even distribution of charge. What must be the direction of the electric field in the bulk of the electrolyte to cause this flow?

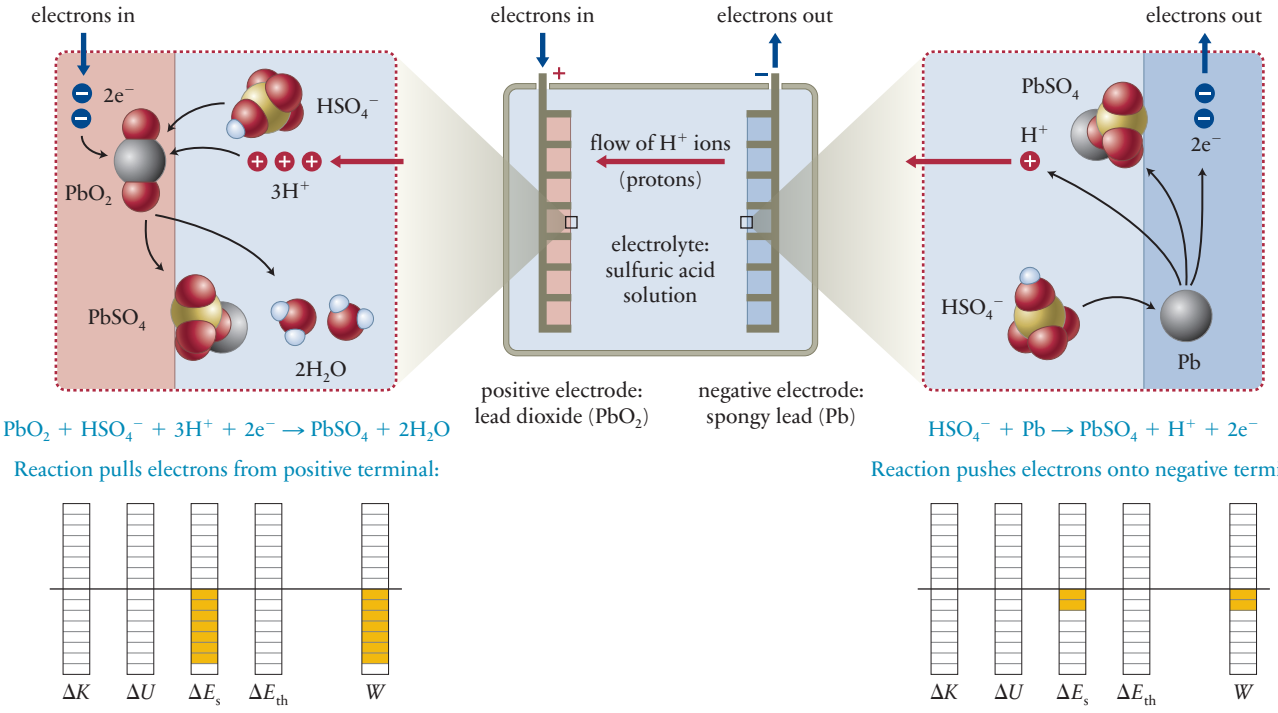
Figure 26.17 illustrates the operation of a lead-acid cell used in automobile batteries. A "12-V" automobile battery consists of six such cells, each producing a potential difference of 2.1 V, for a total of 12.6 V at full charge. The negative electrode of a lead-acid cell is composed of spongy lead (Pb) packed on a metal grid; the positive electrode contains lead dioxide (PbO_2) packed on a metal grid. The electrodes are immersed in sulfuric acid and chemical reactions convert the lead, the lead dioxide, and the sulfuric acid into lead sulfate ($PbSO_4$) and water. For every molecule of lead sulfate that is produced in these reactions, one electron is removed from the positive terminal and one is added to the negative terminal. The left and right sides of Figure 26.17 show energy diagrams for the species undergoing chemical reactions at each of the electrodes. For each reaction, the chemical energy of the species involved in the reaction decreases; this energy is used to do work on the electrons in the electrodes.

26.13 Given that the cell does *positive* work on the electrons, why is it that the work in both energy diagrams in Figure 26.17 is *negative*?

*emf stands for *electromotive force*, a misnomer because this quantity bears no relation to the concept of force. For this reason we shall always refer to this quantity by its abbreviation, rather than its original meaning.

CONCEPTS

Figure 26.17 Schematic diagram of a lead-acid cell and of the reactions taking place at the positive and negative electrodes.



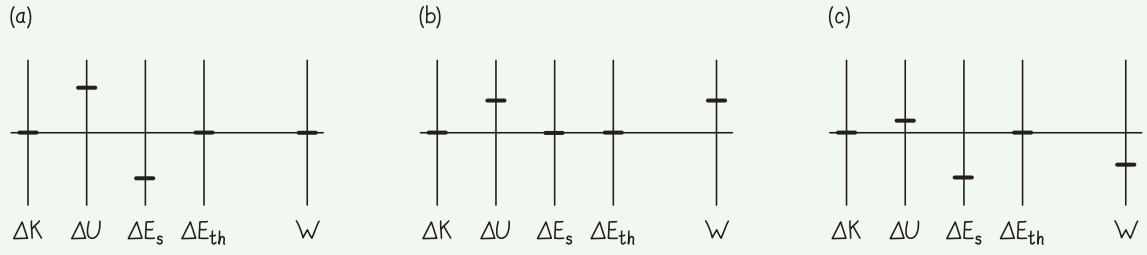
Self-quiz

1. Consider again Figure 26.2 and imagine moving one more electron from the fur to the rod. (*a*) Is the work that must be done on the rod-fur system to accomplish this transfer positive, negative, or zero? (*b*) Is the electrostatic work positive, negative, or zero? (*c*) Does the electric potential energy of the rod-fur system increase, decrease, or remain the same?
2. Figure 22.15 (page 791) shows a person's hair standing out from her head because of "electrostatic charge." Look back at the discussion of Van de Graaff generators and discuss how this can happen when a person touches the globe of the generator but is insulated from the ground.
3. A parallel-plate capacitor is connected to a battery. If the distance between the plates is decreased, do the magnitudes of the following quantities increase, decrease, or stay the same: (*i*) the potential difference between the negative plate and the positive plate, (*ii*) the electric field between the plates, and (*iii*) the charge on the plates?
4. When a dielectric is inserted between the plates of an isolated charged capacitor, do the magnitudes of the following quantities increase, decrease, or stay the same: (*i*) the charge on the plates, (*ii*) the electric field between the plates, and (*iii*) the potential difference between the negative plate and the positive plate?
5. Draw an energy diagram for the process of charging a capacitor with a dielectric as shown in Figure 26.15*b* for the following systems: (*a*) battery, capacitor, and dielectric; (*b*) dielectric only; (*c*) battery and capacitor. Ignore any dissipation of energy.

Answers:

1. (*a*) To displace the electron toward the rod, you must apply a force directed toward the rod. Because the force and force displacement are in the same direction, you (an external agent) must do positive work. (*b*) The electric force exerted on the electron is directed toward the fur, opposite the direction of the force displacement, so the electrostatic work is negative. (*c*) The electric potential energy of the system increases because separating charge carriers increases a system's electric potential energy.
2. If a person is in contact with the globe of the generator but insulated from the ground, then the person acts as an extension of the globe. Electrical charge spreads out over the surface of the person, including the surface of each hair as well. Because each hair has a surplus of the same type of charge, the hairs repel each other and stand out, getting as far away from each other as possible.
3. (*i*) Stays the same. The battery keeps the potential difference across the capacitor constant. (*ii*) To keep a constant potential difference when the distance between the plates decreases, the magnitude of the electric field between the plates must increase because $Ed = V_{\text{batt}}$ (see Example 25.5). (*iii*) For the magnitude of the electric field to increase, the charge on the plates must increase.
4. (*i*) Because the capacitor is isolated, the charge on the plates must remain the same—there is no path for the charge to travel elsewhere. (*ii*) When the dielectric is inserted, the electric field due to the bound surface charge is in the opposite direction of the electric field due to the free charge on the plates and decreases the magnitude of the electric field between the plates. (*iii*) Because the magnitude of the electric field decreases and the separation between the plates is constant, the magnitude of the potential difference between the negative plate and the positive plate must also decrease.
5. See **Figure 26.18**. (*a*) During charging, a decrease in source energy (from the battery) increases the electric potential energy (more charge separation in the dielectric and on the capacitor plates). (*b*) The electric potential energy of the dielectric increases due to work done on it by the battery and the capacitor. The electric potential energy stored in the dielectric is smaller than that stored in part *a* because some electric potential energy is stored on the capacitor plates. (*c*) The decrease in source energy is the same as in part *a*. The electric potential energy stored on the capacitor is smaller than in part *a* because most of the converted source energy ends up in the dielectric, which is not part of the system considered. This energy leaves the system as negative work.

Figure 26.18



26.5 Capacitance

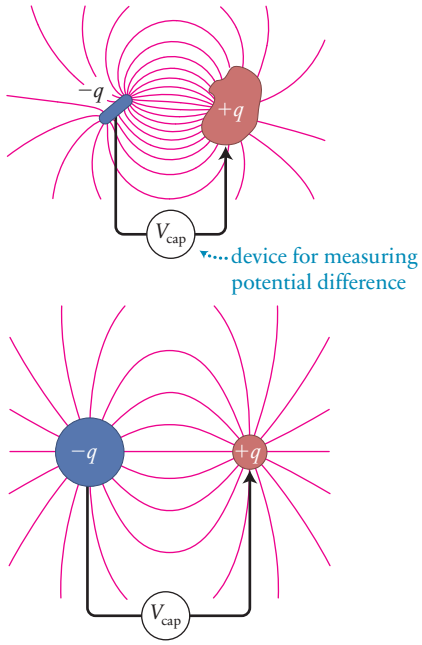
Figure 26.19 shows three capacitors, each one consisting of a pair of conducting objects carrying opposite charges of magnitude q. For each arrangement, the potential difference between the objects is proportional to q; that is, doubling q doubles the potential difference across the capacitor. The ratio of the magnitude of the charge on one of the objects to the magnitude of the potential difference across them is defined as the **capacitance** of the arrangement:

$$C \equiv \frac{q}{V_{\text{cap}}}. \tag{26.1}$$

In Eq. 26.1, q represents *the magnitude of the charge on each conducting object* and V_{cap} is *the magnitude of the potential difference between the conducting objects*. Because both these quantities are positive, C is always positive.

The value of C depends on the size, shape, and separation of the conductors. In Figure 26.19, for example, the values of V_{cap} would typically be different for the three capacitors, even though q is the same for each. Below we'll examine how to determine C for a given set of conductors.

Figure 26.19 The electric fields and potential differences of three different capacitors.



26.14 Two capacitors, A and B, are each connected to a 9-V battery. If $C_A > C_B$, which capacitor stores the greater amount of charge?

The answer to Checkpoint 26.14 suggests a simple interpretation of C. As its name suggests, C represents the capacitor's *capacity to store charge*: The greater C, the greater the amount of charge stored for a given value of V_{cap}.

As you can see from Eq. 26.1, capacitance has SI units of coulomb per volt. This derived unit is given the name **farad** (F), in honor of the English physicist Michael Faraday:

$$1\ \text{F} \equiv 1\ \text{C/V}.$$

As you will see in Checkpoint 26.15, a capacitance of 1 F is enormous. The capacitance of capacitors commonly found in electronic devices is expressed in microfarads ($1\ \mu\text{F} = 1 \times 10^{-6}$ F) and picofarads ($1\ \text{pF} = 1 \times 10^{-12}$ F).

Figure 26.19 suggests a simple procedure for determining the capacitance of a given set of conductors: Determine the potential difference V_{cap} between the two conductors when they carry some given charge q, and use Eq. 26.1 to calculate C. Note that because conductors are equipotentials, V_{cap} represents the potential difference between *any* two points on the conductors measured along *any* path. The Procedure box below gives one procedure for determining the capacitance of a given set of conductors. In the next examples we apply this procedure to some simple configurations of conductors.

PROCEDURE: Calculating the capacitance of a pair of conductors

To calculate the capacitance of a pair of conductors:

1. Let the conductors carry opposite charges of magnitude q.
2. Use Gauss's law, Coulomb's law, or direct integration to determine the electric field along a path leading from the negatively charged conductor to the positively charged conductor.
3. Calculate the electrostatic work W done on a test particle carrying a charge q_t along this path (Eq. 25.24) and determine the potential difference across the capacitor from Eq. 25.15:
$$V_{\text{cap}} = -W_{q_t}(- \rightarrow +)/q_t.$$
4. Use Eq. 26.1, $C \equiv q/V_{\text{cap}}$, to determine C.

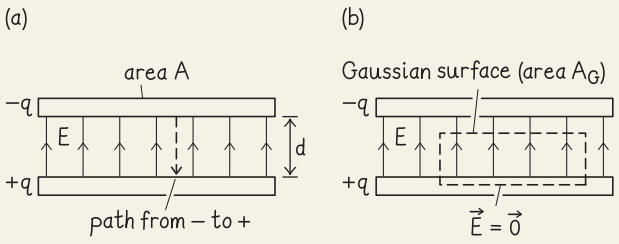
QUANTITATIVE TOOLS

EXAMPLE 26.2 Parallel-plate capacitor

What is the capacitance of a parallel-plate capacitor that has a plate area A and a plate separation distance d?

1 GETTING STARTED I begin by making a sketch of the capacitor, showing the electric field between the plates (**Figure 26.20*a***). The problem doesn't specify the plate shape, so I simply show the capacitor from the side, representing each plate by a horizontal line. If I assume that the separation distance d is small, then the electric field is uniform and confined between the plates.

Figure 26.20



2 DEVISE PLAN I can use the steps of the Procedure box on page 921 to determine the capacitance. The first step is to determine the electric field between the capacitor plates when they carry opposite charges of magnitude q. The second step is to obtain the electrostatic work done on a test particle moved from one plate to the other; I can use Eq. 25.24 to obtain this work. Because the field is uniform, it is most convenient to choose a path along a field line for the path over which the electrostatic work is done. As specified in the Procedure box, the path runs from the negatively charged plate to the positively charged plate. Once I know the electrostatic work, I know the potential difference across the capacitor and so can calculate the capacitance.

3 EXECUTE PLAN Because of the planar symmetry, I can use Gauss's law to determine the electric field. I choose a cylindrical Gaussian surface straddling the surface of the positively charged plate. The cylinder height is less than d, and the area of the end surfaces is A_G (Figure 26.20*b*). The electric flux through this Gaussian surface is zero everywhere except through the top surface, where $\Phi = EA_G$. (The bottom surface is inside the conducting metal plate, where the electric field is zero.)

To apply Gauss's law, I also need to know the charge enclosed by the Gaussian surface. The positive plate carries a charge $+q$ distributed over a surface of area A, so the surface charge density is $\sigma = +q/A$ and the charge enclosed by the Gaussian surface is $q_{enc} = \sigma A_G = (q/A)A_G$. Applying Gauss's law, $\Phi = q_{enc}/\epsilon_0$ (Eq. 24.8), I get

$$EA_G = \frac{q}{\epsilon_0 A} A_G \quad \text{or} \quad E = \frac{q}{\epsilon_0 A} = \frac{\sigma}{\epsilon_0} \tag{1}$$

in agreement with Eq. 24.17.

Now that I know E, I can calculate the electrostatic work required to move a test particle carrying a charge $+q_t$ from the negatively charged to the positively charged plate. The electric force exerted on the test particle is upward in Figure 26.20, and the force displacement is downward because the particle moves from negative plate to positive plate. Because these two vectors point in opposite directions, the electrostatic work done on the test particle is negative, $W_{q_t} = -q_t Ed$, and the potential difference across the capacitor is, from Eq. 25.15,

$$V_{cap} \equiv \frac{-W_{q_t}(- \to +)}{q_t} = \frac{q_t Ed}{q_t} = Ed$$

or, substituting E from Eq. 1,

$$V_{cap} = \frac{qd}{\epsilon_0 A}.$$

Note that the potential difference is proportional to the magnitude of the charge on each plate, q. Using the definition of capacitance, I obtain

$$C \equiv \frac{q}{V_{cap}} = \frac{q}{qd/(\epsilon_0 A)} = \frac{\epsilon_0 A}{d}. \checkmark$$

4 EVALUATE RESULT My result agrees with the conclusions we drew in Section 26.2: The capacitance (or quantity of charge stored for a given potential difference) increases with increasing plate area A and decreasing plate separation distance d. Also, I note that the electric field—and therefore the capacitance—do not depend on the plate: Circular or square plates give the same result.

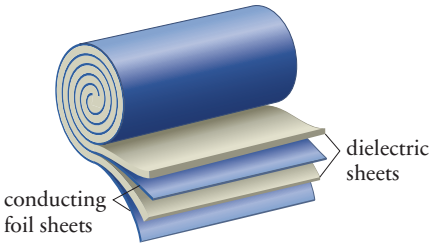
26.15 The plate spacing in a typical parallel-plate capacitor is about 50 μm. (*a*) What is the plate area in a 1-μF capacitor? (*b*) Given that the electric field at which electrical breakdown occurs in air is about 3×10^6 V/m, what is the maximum charge that this capacitor can hold? (*c*) How many electrons does this charge correspond to? (The electron's charge is $e = 1.6 \times 10^{-19}$ C.)

As Checkpoint 26.15 shows, even modest capacitances require very large plate areas. Various techniques are used to keep the overall size of capacitors small, one of which involves rolling up two thin conducting sheets that are separated by thin sheets of a dielectric material (**Figure 26.21*a***). Figure 26.21*b* shows a number of different capacitors used in electronic circuits.

QUANTITATIVE TOOLS

Figure 26.21

(*a*) One way to design a compact capacitor with a large surface area



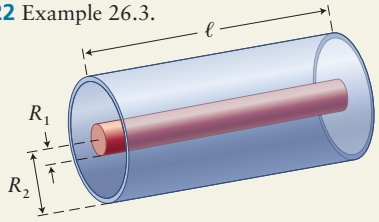
(*b*) Some capacitors used in electronic circuits



EXAMPLE 26.3 Coaxial cylindrical capacitor

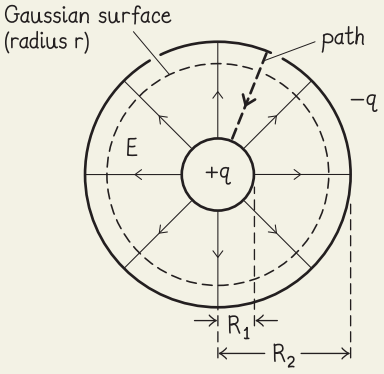
Figure 26.22 shows a *coaxial capacitor* consisting of two concentric metal cylinders 1 and 2, of radii R_1 and $R_2 > R_1$, and both of length $\ell \gg R_2$. Both cylinders are made of metal. What is the capacitance of this arrangement?

Figure 26.22 Example 26.3.



❶ **GETTING STARTED** To determine the capacitance, I must let the two cylinders carry opposite charges of magnitude q, which I assume to be uniformly distributed over each cylinder. If I let cylinder 1 carry a charge $+q$ and cylinder 2 carry a charge $-q$, the electric field points radially outward from cylinder 1 to cylinder 2 (**Figure 26.23**). Because the cylinders are very long relative to their separation distance $R_2 - R_1$, I assume that the electric field is confined to the volume between the cylinders.

Figure 26.23



❷ **DEVISE PLAN** Again I refer to the Procedure box on page 921 to calculate the capacitance. For the path over which electrostatic work is done, I choose a straight path that goes radially from cylinder 2 to cylinder 1.

❸ **EXECUTE PLAN** Because of the cylindrical symmetry, I choose a cylindrical Gaussian surface (Figure 26.23). The length of the Gaussian surface is ℓ_G, and its radius is r ($R_2 > r > R_1$). The electric flux Φ through the curved portion of the Gaussian surface is equal to the product of the electric field strength E_r at a distance r from the common axis of cylinders 1 and 2 and the surface area of the Gaussian surface $A_G = (2\pi r)\ell_G$. Therefore $\Phi = 2\pi r \ell_G E_r$. Because the linear charge density on cylinder 1 is $+q/\ell$, the quantity of charge enclosed by the Gaussian surface is given by the product of the linear charge density and the length of the Gaussian surface: $q_{enc} = +(q/\ell)\ell_G$. Applying Gauss's law, I get

$$2\pi r \ell_G E_r = \frac{q\ell_G}{\epsilon_0 \ell},$$

or $E_r = q/(2\pi\epsilon_0 \ell r)$, in agreement with the result we obtained in Exercise 24.7, $E = 2k\lambda/r$, because $k = 1/(4\pi\epsilon_0)$ and $q/\ell = \lambda$.

Now that I know E_r, I can calculate the electrostatic work required to move a test particle carrying a charge q_t from cylinder 2 to cylinder 1. Integrating the electric force exerted on the test particle over the force displacement from cylinder 2 (negatively charged) to cylinder 1 (positively charged), I get

$$W_{q_t} = \int_{R_2}^{R_1} \frac{qq_t}{2\pi\epsilon_0 \ell r}\, dr.$$

Working out the integral, I obtain

$$W_{q_t} = \frac{qq_t}{2\pi\epsilon_0 \ell}\Big[\ln r\Big]_{R_2}^{R_1} = \frac{qq_t}{2\pi\epsilon_0 \ell}\ln\left(\frac{R_1}{R_2}\right).$$

The potential difference between the negative cylinder 2 and the positive cylinder 1 is thus

$$V_{cap} \equiv \frac{-W_{q_t}}{q_t} = -\frac{q}{2\pi\epsilon_0 \ell}\ln\left(\frac{R_1}{R_2}\right) = \frac{q}{2\pi\epsilon_0 \ell}\ln\left(\frac{R_2}{R_1}\right).$$

Because $R_2 > R_1$, the logarithm is positive and therefore V_{cap} is positive, as it should be because I am bringing a quantity q_t of positive charge from a location of low potential on negatively charged cylinder 2 to a location of high potential on positively charged cylinder 1.

(Continued)

According to Eq. 26.1, the capacitance of the coaxial capacitor is thus

$$C \equiv \frac{q}{V_{\text{cap}}} = \frac{2\pi\epsilon_0\ell}{\ln(R_2/R_1)}. \checkmark$$

❹ **EVALUATE RESULT** My result shows that the capacitance is proportional to ℓ, which makes sense: The longer the coaxial cylinders, the greater the quantity of charge that can be stored on them. Decreasing R_1 or increasing R_2 is equivalent to increasing the plate separation distance d in a parallel-plate capacitor, which decreases the capacitance. Indeed, my result shows a decreasing capacitance for decreasing R_1 or increasing R_2. (The dependence on R_1 and R_2 is a bit more complicated than the dependence on d in a parallel-plate capacitor because the electric field in the coaxial capacitor is nonuniform and because changing the radii of the cylinders affects their surface areas.)

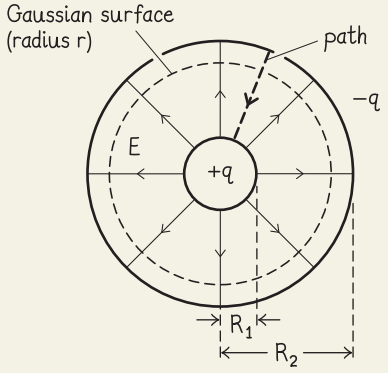
26.16 Coaxial cables used for cable television typically have a central metallic core of 0.20-mm radius, surrounded by a cylindrical metallic sheath of 2.0-mm radius. The two are separated by a plastic spacer. If the effect of the spacer can be ignored (that is, assuming the two conductors are separated by air), what is the capacitance of a 100-m-long cable?

EXAMPLE 26.4 Spherical capacitor

What is the capacitance of a spherical capacitor consisting of two concentric conducting spherical shells of radii R_1 and $R_2 > R_1$?

❶ **GETTING STARTED** If I let the inner sphere carry a positive charge $+q$ and the outer one a negative charge $-q$, my sketch of the capacitor looks identical to the sketch of the coaxial capacitor of Example 26.3 (**Figure 26.24**). The calculation, however, will not be the same because now the electric field has spherical, not cylindrical, symmetry.

Figure 26.24



❷ **DEVISE PLAN** For the path over which electrostatic work is done, I choose a straight path that goes radially from the outer sphere to the inner sphere. The outer sphere does not contribute to the electric field between the spheres because the field inside a hollow conductor is always zero (see Section 24.5). The electric field created by the inner sphere is the same as that created by a charged particle (Eq. 24.15, $E = kq/r^2$), so I can use this expression to follow steps 3 and 4 in the Procedure box on page 921.

❸ **EXECUTE PLAN** The electrostatic work done in moving a test particle carrying a charge q_t from the outer sphere to the inner sphere is

$$W_{q_t}(- \to +) = \int_{R_2}^{R_1} \frac{qq_t}{4\pi\epsilon_0 r^2}\, dr.$$

Working out the integral, I get

$$W_{q_t}(- \to +) = -\frac{qq_t}{4\pi\epsilon_0}\left[\frac{1}{r}\right]_{R_2}^{R_1} = -\frac{qq_t}{4\pi\epsilon_0}\left[\frac{1}{R_1} - \frac{1}{R_2}\right].$$

The potential difference between the outer and inner spheres is thus

$$V_{\text{cap}} \equiv \frac{-W_{q_t}(- \to +)}{q_t} = \frac{q}{4\pi\epsilon_0}\left[\frac{1}{R_1} - \frac{1}{R_2}\right],$$

so the capacitance is

$$C \equiv \frac{q}{V_{\text{cap}}} = 4\pi\epsilon_0\left[\frac{1}{R_1} - \frac{1}{R_2}\right]^{-1} = 4\pi\epsilon_0\left[\frac{R_2 - R_1}{R_1 R_2}\right]^{-1}$$

$$= 4\pi\epsilon_0\left[\frac{R_1 R_2}{R_2 - R_1}\right]. \checkmark$$

❹ **EVALUATE RESULT** I expect the capacitance to go up as the separation distance $R_2 - R_1$ between the spheres decreases, and this is just what my result shows. If I increase the spheres' radii while keeping their separation distance $R_2 - R_1$ fixed, the surface area $A = 4\pi R^2$ of each sphere increases and the capacitance should increase, in agreement with my result.

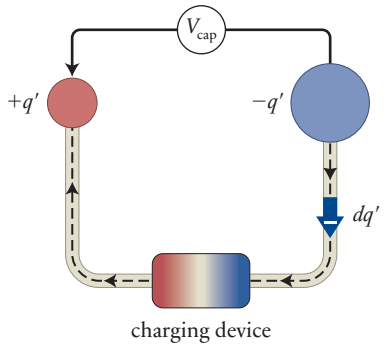
26.17 (*a*) To calculate the "capacitance" of an isolated sphere, evaluate the result we obtained in Example 26.4 in the limit that R_2 goes to infinity. (*b*) What is the capacitance of the spherical metal dome of a Van de Graaff generator that has a radius of about 2.5 m. (*c*) Given that air breaks down in an electric field with a magnitude of about 3.0×10^6 V/m, what is the maximum amount of charge that can be stored on the dome before the air breaks down?

QUANTITATIVE TOOLS

26.6 Electric field energy and emf

How much electric potential energy is stored in a charged capacitor? To answer this question, consider a simple capacitor consisting of two conducting objects. In order to charge the capacitor, some charge-separating device must transfer charge from one conductor to the other (Figure 26.25). During this transfer, the charge-separating device does work on the capacitor and this energy ends up as electric potential energy "stored in the capacitor."* One complication in the calculation of the work done by the charge-separating device is that as the magnitude of the charge on each conductor increases, the potential difference increases too, so the work required to transfer a unit of charge increases.

Figure 26.25 To determine the electric potential energy stored in a capacitor, we calculate the energy required to transfer charge from one conductor to the other.



Let us therefore break down the transfer of charge from one conductor to the other into small increments of charge dq', so that the potential difference is essentially constant during the transfer of a single increment. Consider some instant during the charging when the magnitude of the charge on each conductor is q'. The potential difference between the negative and positive conductors is then given by Eq. 26.1: $V_{cap} = q'/C$. As an additional increment of charge dq' is moved from the negative to the positive conductor, the electrostatic work done on it is $dW = -dq'V_{cap}$ (Eq. 25.17). Because the charge-separating device must do work on charge carriers against the electric force, the work done by the charge-separating device on the charge carriers is the negative of the electrostatic work, so the change in electric potential energy of the capacitor during the transfer is

$$dU^E = -dW = V_{cap}\,dq' = \frac{q'}{C}\,dq'. \tag{26.2}$$

When the magnitude of the charge on each conductor has increased from zero to its final value q, the electric potential energy stored in the capacitor is

$$U^E = \int dU^E = \int_0^q \frac{q'}{C}\,dq' = \frac{1}{C}\int_0^q q'\,dq' = \tfrac{1}{2}\frac{q^2}{C}. \tag{26.3}$$

Often it is more convenient to express the electric potential energy not in terms of the magnitude of the charge q on the capacitor, but in terms of the potential difference across it:

$$U^E = \tfrac{1}{2}\frac{q^2}{C} = \tfrac{1}{2}CV_{cap}^2 = \tfrac{1}{2}qV_{cap}. \tag{26.4}$$

Note that Eqs. 26.3 and 26.4 hold for any type of capacitor, regardless of the configuration of the conductors. All that enters into these expressions besides the charge or the potential difference is the capacitance, which depends on the size, shape, and the separation of the conductors.

26.18 A 1.0-μF parallel-plate capacitor with a plate spacing of 50 μm is charged up to the breakdown threshold. (*a*) If the electric field in the air between the capacitor plates is 3.0×10^6 V/m, how much energy is stored in the capacitor? Express your answer in joules. (*b*) How high must you raise this book ($m \approx 2$ kg) to increase the gravitational potential energy of the Earth-book system by the same amount?

As we discussed in Section 26.1, we can imagine electric potential energy to be stored either in the configuration of charge in the capacitor or in the electric field. We can use Eq. 26.4 and our knowledge about the

*We are assuming that there is no dissipation of energy, so that all the work done on the system ends up as electric potential energy. In practice this is a reasonable assumption.

QUANTITATIVE TOOLS

electric field in a capacitor to relate electric potential energy to the electric field. From Example 25.5 we know that the magnitude of the potential difference between the plates of a parallel-plate capacitor is given by Ed, so, using the expression for C in Example 26.2 and Eq. 26.4, we can write for the electric potential energy stored in a parallel-plate capacitor

$$U^E = \tfrac{1}{2} C V_{\text{cap}}^2 = \tfrac{1}{2}\left(\frac{\epsilon_0 A}{d}\right)(Ed)^2 = \tfrac{1}{2}\epsilon_0 E^2 (Ad). \tag{26.5}$$

The term in parentheses on the right side, Ad, is equal to the volume of the space between the capacitor plates—that is, the region to which the electric field is confined. Therefore, the energy per unit volume stored in the electric field—the **energy density** of the electric field—is

$$u_E \equiv \frac{U^E}{\text{volume}} = \tfrac{1}{2}\epsilon_0 E^2. \tag{26.6}$$

Although we derived this expression for the special case of a parallel-plate capacitor, it holds true for any electric field in vacuum. Any given region of space where a uniform electric field is present can be viewed as containing an amount of electric potential energy equal to $\frac{1}{2}\epsilon_0$ times the square of the magnitude of the electric field in that region times the volume. If the electric field is nonuniform, we must subdivide the volume of interest into small enough segments that E can be considered uniform within each segment, then apply Eq. 26.6 to each segment and take the sum of all the contributions. (This corresponds to integrating the energy density over the volume of the region that contains the electric field.)



26.19 A parallel-plate capacitor has plates of area A separated by a distance d. The magnitude of the charge on each plate is q. (*a*) Determine the magnitude of the force exerted by the positively charged plate on the negatively charged one. (*b*) Suppose you increase the separation between the plates by an amount Δx. How much work do you need to do on the capacitor to achieve this increase? (*c*) What is the change in the electric potential energy of the capacitor? (*d*) Moving the plate adds additional space with electric field between the plates. Show that the energy stored in the electric field in this additional space is equal to the work done on the capacitor.

The energy stored in a capacitor is supplied to it by the charging device—such as a generator, a battery, or a solar cell. Inside this device, nonelectrostatic interactions cause a separation of charge by doing work on charged particles. The work per unit charge done by the nonelectrostatic interactions on the charge carriers inside the device is called the emf and is denoted by $\mathscr{E}$:

$$\mathscr{E} \equiv \frac{W_{\text{nonelectrostatic}}}{q}. \tag{26.7}$$

The SI unit of emf is the same as that of potential: the volt. The rating of a battery—1.5 V or 9 V—gives its emf.*

If no energy is dissipated inside the charging device, *all* of the energy can be transferred to charge carriers outside the device. This transfer takes

*The term *voltage* is sometimes used to refer to a potential difference or to an emf (such as the rating of a battery). Potential difference, however, is related to *electrostatic* work done on charge carriers, whereas emf deals with *nonelectrostatic* work done on them. Thus, electrostatic work brings opposite charge carriers together, while nonelectrostatic work causes charge separation. To maintain this important distinction, we shall avoid the term *voltage*.

QUANTITATIVE TOOLS

place through electric interactions. In Figure 26.25, for example, electric forces remove electrons from one object and push them onto the other, charging the capacitor. In the absence of any energy dissipation, the nonelectrostatic work done on charge carriers inside the device is equal to the electrostatic work done on charge carriers outside it. Because the electrostatic work per unit charge is the potential difference between the negative and positive terminals of the charging device, we have, for an ideal charging device,

$$V_{\text{device}} = \mathscr{E} \quad \text{(ideal device)}. \tag{26.8}$$

In practice, some energy is always dissipated inside the device, so not all of the nonelectrostatic work done on charge carriers inside the device can be turned into electrostatic work. Consequently, for most devices, $V_{\text{device}} < \mathscr{E}$.

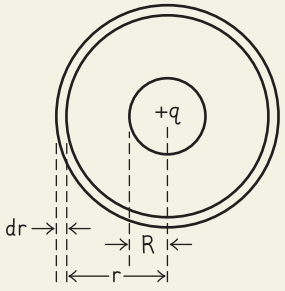
EXAMPLE 26.5 Van de Graaff energy

The radius of the dome on one very large Van de Graaff generator is about 2.5 m, and air breaks down when the field magnitude is about 3.0×10^6 V/m. How much electric potential energy is stored in the electric field surrounding the dome just before the air there breaks down?

❶ **GETTING STARTED** I am given the radius of a Van de Graaff dome and asked to calculate how much potential energy is in the electric field surrounding the dome just before the field causes the air to break down. If I approximate the dome as a uniformly charged spherical shell, then the electric field surrounding it is the same as that surrounding a particle carrying the same charge [Eq. 24.15, $E = q/(4\pi\epsilon_0 r^2)$]. The magnitude of the electric field around the dome is greatest at the dome surface, so I take this E_{surf} value as the maximum value just before the air breaks down.

❷ **DEVISE PLAN** Equation 26.6 gives me the energy density of the electric field around the dome. I can substitute the Eq. 24.15 expression for the electric field of a sphere carrying a charge q into Eq. 26.6 to obtain an expression for the energy density of the electric field at an arbitrary distance r from the dome center. Because the electric field has the same magnitude at any location a distance r from the dome center, I can divide the space outside the dome into a series of thin spherical shells, each of thickness dr and all concentric with the dome (**Figure 26.26**), and then integrate over all shells from $r = R$ to $r = \infty$ to obtain the energy stored in the electric field surrounding the dome in terms of the charge q. I can then use Eq. 24.15 to eliminate q from my result and express the energy stored in terms of E_{surf}, which is given.

Figure 26.26



❸ **EXECUTE PLAN** Substituting Eq. 24.15 into Eq. 26.6, I get

$$u_E = \tfrac{1}{2}\epsilon_0 E^2 = \tfrac{1}{2}\epsilon_0\left(\frac{q}{4\pi\epsilon_0 r^2}\right)^2.$$

The volume of a thin spherical shell of radius r and thickness dr centered on the dome is equal to the surface area of a sphere of radius r times the shell thickness: $(4\pi r^2)dr$. The energy in that volume is thus $dU^E = u_E(4\pi r^2)dr$, and so the electric potential energy in the space around the dome is

$$\begin{aligned} U^E &= \int_R^\infty dU^E = \int_R^\infty u_E(4\pi r^2)\,dr \\ &= \tfrac{1}{2}\epsilon_0\int_R^\infty\left(\frac{q}{4\pi\epsilon_0 r^2}\right)^2(4\pi r^2)dr = \frac{q^2}{8\pi\epsilon_0}\int_R^\infty \frac{1}{r^2}\,dr \\ &= \frac{q^2}{8\pi\epsilon_0}\left[\frac{-1}{r}\right]_R^\infty = \frac{q^2}{8\pi\epsilon_0 R}. \end{aligned} \tag{1}$$

Now I have an expression for U^E, the quantity I must determine, but I have no value for q. Given that the electric field at the dome surface is given by Eq. 24.15, $E_{\text{surf}} = q/(4\pi\epsilon_0 R^2)$, I can rearrange this expression to $q = E_{\text{surf}}(4\pi\epsilon_0 R^2)$ and rewrite Eq. 1 as $U^E = 2\pi\epsilon_0 E_{\text{surf}}^2 R^3$. Because E_{surf} is the maximum electric field magnitude around the dome, I know that this magnitude must be the breakdown value for air. Substituting the values given, I get

$$\begin{aligned} U^E &= 2\pi[8.85\times10^{-12}\,\text{C}^2/(\text{N}\cdot\text{m}^2)](3.0\times10^6\,\text{V/m})^2(2.5\,\text{m})^3 \\ &= 7.8\,\text{kJ}. \checkmark \end{aligned}$$

❹ **EVALUATE RESULT** As a check on my work, I can calculate the electric potential energy of the charge stored on the dome using Eq. 26.4. In Checkpoint 26.17, you found that the capacitance of an isolated sphere is $C_{\text{sphere}} = 4\pi\epsilon_0 R$ and that the potential of a sphere is related to the electric field at its surface by $V_{\text{cap}} = ER$. Therefore $U^E = \frac{1}{2}CV_{\text{cap}}^2 = 2\pi\epsilon_0 R^3 E_{\text{surf}}^2$, which is the same result I obtained.

QUANTITATIVE TOOLS

26.20 To store electric potential energy, the flash unit on a typical camera uses a 100-μF capacitor charged to a potential difference of 300 V. When the flash is fired, the energy in the capacitor is released to a bulb in a burst of about 1.0-ms duration. (*a*) How much energy is stored in the fully charged capacitor before it is fired? (*b*) What is the average power of the flash firing?

26.7 Dielectric constant

As we saw in Section 26.3, the capacitance of a capacitor can be increased by inserting a dielectric between the two conductors. For example, inserting a slab of mica between the plates of an isolated charged capacitor (**Figure 26.27**) decreases the potential difference across the capacitor by a factor of 5. This tells us that the mica reduces the electric field inside the isolated capacitor by a factor of 5. By definition, the magnitude of the potential difference V_0 across the isolated capacitor without a dielectric divided by the magnitude of the potential difference V_d with the dielectric is called the **dielectric constant** κ:

$$\kappa \equiv \frac{V_0}{V_d}. \tag{26.9}$$

Given that the magnitude q_0 of the charge on each plate of the isolated capacitor is not affected by the dielectric, we see from Eqs. 26.9 and 26.1 that

$$\kappa \equiv \frac{V_0}{V_d} = \frac{V_0/q_0}{V_d/q_0} = \frac{1/C_0}{1/C_d} = \frac{C_d}{C_0}, \tag{26.10}$$

where C_d is the capacitance of the capacitor with the dielectric and C_0 that without a dielectric. Therefore, the capacitance changes by the factor κ when a dielectric is inserted:

$$C_d = \kappa C_0. \tag{26.11}$$

The dielectric constant is always greater than or equal to 1 ($\kappa \geq 1$). It is greater than 1 with a dielectric in place because the presence of a dielectric decreases the electric field inside the capacitor. The greater the polarization of the dielectric material, the more reduced the electric field inside the dielectric and the greater the dielectric constant κ. **Table 26.1** gives the dielectric constants for several commonly used dielectric materials.

Figure 26.27 The potential difference across an isolated parallel-plate capacitor is greater (*a*) without a dielectric between the plates than it is (*b*) with a dielectric.

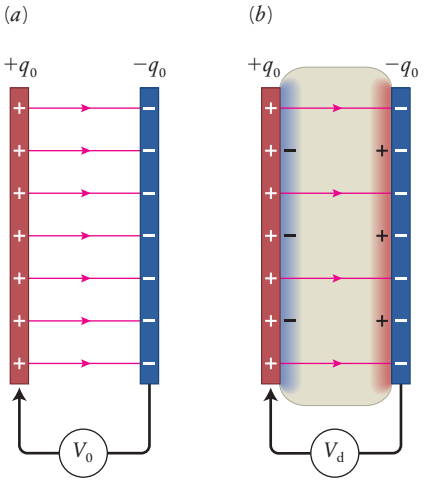


Table 26.1 Dielectric properties

Material	Dielectric constant κ	Breakdown threshold E_{max} (V/m)
Air (1 atm)	1.00059	3.0×10^6
Paper	1.5–3	4.0×10^7
Mylar (polyester)	3.3	4.3×10^8
Quartz	4.3	8×10^6
Mica	5	2×10^8
Oil	2.2–2.7	
Porcelain	6–8	
Water (distilled, 20°C)	80.2	$6.5–7 \times 10^7$
Titania ceramic	126	8×10^6
Strontium titanate	322	
Barium titanate	1200	8×10^7

The dielectric constant for vacuum—that is, no material between the plates—is unity (one) by definition. Because air is very dilute, the dielectric constant of air is nearly unity as well. If the dielectric is composed of polar molecules that can align themselves (such as the water molecules in liquid water), then the overall polarization is much greater than in nonpolar dielectrics, so the dielectric constant is large. For some polar materials, the dielectric constant can be in the thousands.

The electric field $\vec{E}$ inside the dielectric is the superposition of the electric field due to the free charge on the plates, $\vec{E}_{\text{free}}$, and the electric field due to the bound surface charge on the dielectric, $\vec{E}_{\text{bound}}$: $\vec{E} = \vec{E}_{\text{free}} + \vec{E}_{\text{bound}}$ (**Figure 26.28a**). We designate the magnitude of the free charge on the capacitor plates by q_{free} and the magnitude of the bound charge on the surfaces of the dielectric by q_{bound} (Figures 26.28*b* and *c*). With this notation, both q_{free} and q_{bound} are always positive. Using the expression for the electric field of a sheet of charge, we can thus write for the magnitude of the electric field inside the capacitor in the absence of a dielectric

$$E_{\text{free}} = \frac{\sigma_{\text{free}}}{\epsilon_0} = \frac{q_{\text{free}}}{\epsilon_0 A}, \tag{26.12}$$

where $\sigma_{\text{free}} = q_{\text{free}}/A$ is the magnitude of the free surface charge density and A is the area of either capacitor plate. Likewise, the magnitude of the electric field due to the bound surface charge is

$$E_{\text{bound}} = \frac{\sigma_{\text{bound}}}{\epsilon_0} = \frac{q_{\text{bound}}}{\epsilon_0 A}, \tag{26.13}$$

where σ_{bound} is the magnitude of the bound surface charge density. Because $\vec{E}_{\text{free}}$ and $\vec{E}_{\text{bound}}$ point in opposite directions, the magnitude of the electric field $\vec{E}$ inside the dielectric is then

$$E = E_{\text{free}} - E_{\text{bound}} = \frac{\sigma_{\text{free}} - \sigma_{\text{bound}}}{\epsilon_0} = \frac{q_{\text{free}} - q_{\text{bound}}}{\epsilon_0 A}. \tag{26.14}$$

Let us determine the magnitude of the bound charge q_{bound}. If the plate separation is d, we can write Eq. 26.10 in the form

$$\kappa \equiv \frac{V_0}{V_d} = \frac{E_{\text{free}}\, d}{E\, d} = \frac{E_{\text{free}}}{E}, \tag{26.15}$$

where E_{free} is the magnitude of the electric field due to the free charge only, and E is the magnitude of the electric field inside the dielectric. In other words, the dielectric reduces the electric field by the factor κ: $E = E_{\text{free}}/\kappa$. Substituting Eqs. 26.12 and 26.14 into this expression, we get

$$\frac{q_{\text{free}}}{\kappa \epsilon_0 A} = \frac{q_{\text{free}} - q_{\text{bound}}}{\epsilon_0 A} \tag{26.16}$$

or

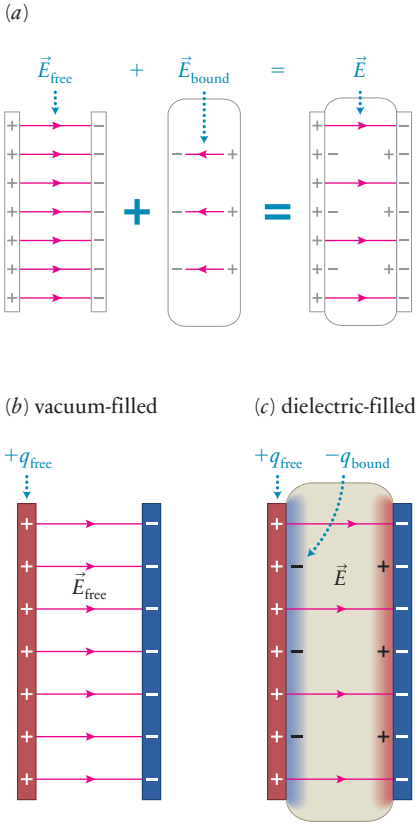
$$\frac{q_{\text{free}}}{\kappa} = q_{\text{free}} - q_{\text{bound}}. \tag{26.17}$$

Solving this expression for q_{bound}:

$$q_{\text{bound}} = \frac{\kappa - 1}{\kappa} q_{\text{free}}. \tag{26.18}$$

Because κ is always greater than 1 in the presence of a dielectric, we see that the magnitude of the bound surface charge is always smaller than the magnitude of the free charge that causes it.

Figure 26.28 (*a*) The electric field inside a dielectric-filled capacitor is the vector sum of the electric field due to the charged plates and that due to the polarized dielectric. (*b*) and (*c*) Bound and free charge on a vacuum-filled and a dielectric-filled isolated parallel-plate capacitor.



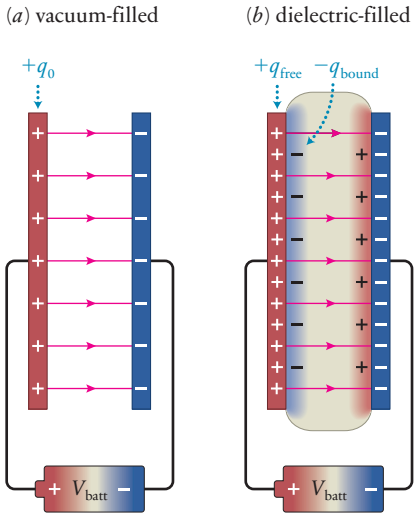


Figure 26.29 Bound and free charge on a vacuum-filled and a dielectric-filled parallel-plate capacitor connected to a battery.

Next we consider the situation of a capacitor connected to a battery (**Figure 26.29**). In this situation, the potential difference across the capacitor is constant, but the charge on the plates changes when the dielectric is inserted. As we have seen in Section 26.3, the electric field inside the dielectric must be the same as before the dielectric was inserted. In other words, the sum of the free and bound charges must still be equal to q_0; that is, $q_0 = q_{free} - q_{bound}$ (Figure 26.29). Because the definition of capacitance involves only the charge on the capacitor plates, we can write

$$C_d \equiv \frac{q_{free}}{V_d} \tag{26.19}$$

and likewise

$$C_0 \equiv \frac{q_0}{V_0}. \tag{26.20}$$

Note the difference between q_{free} and q_0. Even though both represent free charge, they are not equal because when the dielectric is inserted, the battery increases the charge on the plate in order to maintain a constant potential difference across the capacitor, so $q_{free} > q_0$. Because $C_d = \kappa C_0$ (Eq. 26.11), we see from Eqs. 26.19 and 26.20 that the dielectric increases the charge on the capacitor plates by the factor κ:

$$q_{free} = \kappa q_0. \tag{26.21}$$

26.21 (*a*) In Figure 26.29, what is the magnitude of q_{bound}? Express your answer in terms of q_0 and the properties of the dielectric. (*b*) What is the bound surface charge density on the dielectric? Express your answer in terms of the electric field E.

EXAMPLE 26.6 Capacitor with dielectric

A parallel-plate capacitor consists of two conducting plates with a surface area of 1.0 m^2 and a plate separation distance of 50 μm. (*a*) Determine the capacitance and the energy stored in the capacitor when it is charged by connecting it to a 9.0-V battery. (*b*) With the capacitor fully charged and disconnected from the battery, a 50-μm-thick sheet of Mylar is inserted between the plates. Determine the potential difference across the capacitor and the energy stored in it. (*c*) If the Mylar-filled capacitor is connected to the battery, how much work does the battery do to fully charge the capacitor?

❶ **GETTING STARTED** I am given information about a capacitor and a battery used to charge it. From this information, I must determine the capacitance and the energy stored in the capacitor with and without a sheet of Mylar between the plates connected to the battery, and determine what happens to the potential with and without a sheet of Mylar between the plates and with and without the battery connected to the capacitor. When connected to the capacitor, the battery maintains a constant potential across the capacitor. When the battery is not connected to the capacitor, the charge on the plates remains constant.

❷ **DEVISE PLAN** To calculate the energy stored in the capacitor I can use Eq. 26.4; I calculated the capacitance of a parallel-plate capacitor in Example 26.2. When the dielectric is added, the potential and the capacitance are given by Eqs. 26.9 and 26.11. From Table 26.1, I see that the dielectric constant of Mylar is $\kappa = 3.3$.

❸ **EXECUTE PLAN** (*a*) Using the result of Example 26.2, I get

$$C = \frac{[8.85 \times 10^{-12}\ \mathrm{C^2/(N \cdot m^2)}](1.0\ \mathrm{m^2})}{50 \times 10^{-6}\ \mathrm{m}}$$

$$= 0.18 \times 10^{-6} \frac{\mathrm{C^2}}{\mathrm{N \cdot m}} = 0.18\ \mu\mathrm{F}, \checkmark$$

and Eq. 26.4 gives

$$U^E = \tfrac{1}{2} C(V_0)^2 = \tfrac{1}{2} (0.18\ \mu\mathrm{F})(9.0\ \mathrm{V})^2 = 7.2\ \mu\mathrm{J}. \checkmark$$

(*b*) Because of the bound surface charge on the dielectric, the electric field between the capacitor plates decreases, and so the potential difference across the capacitor decreases, too. From the definition of the dielectric constant (Eq. 26.9), I have

$$V_d = \frac{V_0}{\kappa} = \frac{9.0\ \mathrm{V}}{3.3} = 2.7\ \mathrm{V}, \checkmark$$

QUANTITATIVE TOOLS

where I obtained my value for κ from Table 26.1. To calculate the energy in the presence of the dielectric, I must first obtain an expression for the capacitance of the dielectric-filled capacitor. Substituting the expression for the capacitance of a parallel-plate capacitor (see Example 26.2) into Eq. 26.11 yields

$$C_d = \kappa C_0 = \frac{\kappa \epsilon_0 A}{d}$$

$$= \frac{(3.3)[8.85 \times 10^{-12}\ \mathrm{C^2/(N \cdot m^2)}](1.0\ \mathrm{m^2})}{50 \times 10^{-6}\ \mathrm{m}}$$

$$= 0.58\ \mu\mathrm{F}.$$

The stored energy is thus

$$U^E = \tfrac{1}{2} C V_d^2 = \tfrac{1}{2} (0.58\ \mu\mathrm{F})(2.7\ \mathrm{V})^2 = 2.1\ \mu\mathrm{J}. \checkmark$$

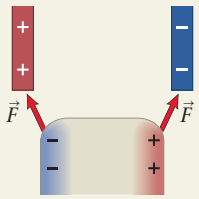
(*c*) The energy stored in the fully charged dielectric-filled capacitor is

$$U^E = \tfrac{1}{2} C V_{batt}^2 = \tfrac{1}{2} (0.58\ \mu\mathrm{F})(9.0\ \mathrm{V})^2 = 24\ \mu\mathrm{J}.$$

From part *b* I know that before it was connected to the battery, the capacitor stored 2.1 μJ, and so the work done by the battery in charging the capacitor must be 24 μJ − 2.1 μJ = 22 μJ. ✔

4 EVALUATE RESULT My answers to parts *a* and *b* show that the amount of energy stored decreases when the dielectric is inserted. That makes sense because, as the dielectric is brought near the plates, the charged plates induce a polarization on the dielectric and consequently it is pulled into the space between the plates (**Figure 26.30**). Therefore, the capacitor does *positive* work on whoever is holding the dielectric, and the energy in the capacitor decreases as the dielectric enters the space between the plates. This work is equal to the difference in energy between parts *a* and *b*: $W = 7.2\ \mu\mathrm{J} - 2.1\ \mu\mathrm{J} = 5.1\ \mu\mathrm{J}$. My answer to part *c* is about three times greater than the value I calculated for U^E in part *a*, which is what I expect given that the capacitance is increased by the factor $\kappa = 3.3$ once the dielectric is inserted.

Figure 26.30



26.22 Verify that in the solution to part *a* of Example 26.6, (*a*) the ratio of units $\mathrm{C^2/(N \cdot m)}$ is equivalent to the unit F and (*b*) the product of units $\mathrm{F \cdot V^2}$ is equivalent to the unit J.

26.8 Gauss's law in dielectrics

Can we apply Gauss's law to calculate the electric fields inside dielectric materials? The answer is *yes*, because Gauss's law is a fundamental law that follows directly from the $1/r^2$ dependence of Coulomb's law. Thus, the presence of a dielectric cannot affect its validity.

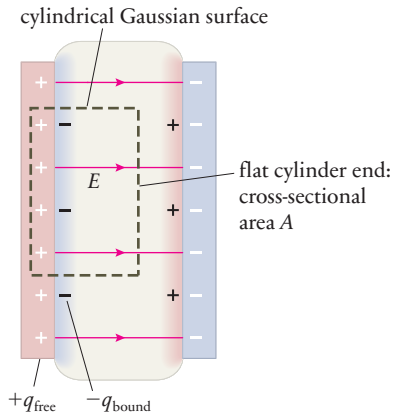
Consider the situation illustrated in **Figure 26.31**. To determine the magnitude of the electric field E inside the dielectric, we consider the cylindrical Gaussian surface with cross-sectional area A shown in the figure. The electric flux is zero except through the right flat surface of the cylinder, so

$$\oint \vec{E} \cdot d\vec{A} = EA. \qquad (26.22)$$

The charge enclosed by the Gaussian surface is not just the enclosed charge on the plate—we must also take into account the enclosed bound charge on the dielectric. The enclosed charge is thus $q_{enc} = q_{free,\,enc} - q_{bound,\,enc}$, and Gauss's law then gives

$$\oint \vec{E} \cdot d\vec{A} = EA = \frac{q_{free,\,enc} - q_{bound,\,enc}}{\epsilon_0}. \qquad (26.23)$$

Figure 26.31 A cylindrical Gaussian surface used to calculate the electric field inside a dielectric-filled parallel-plate capacitor.



QUANTITATIVE TOOLS

In this form, Gauss's law is not very useful, because in order to extract E from Eq. 26.23, we need to know the magnitude of the bound surface charge. Generally, we don't know the contribution from the bound charge in a given situation.

Substituting the relationship between the free and bound charges (Eq. 26.18), however, we can rewrite Eq. 26.23 in the form

$$\oint \vec{E} \cdot d\vec{A} = \frac{q_{\text{free, enc}}}{\epsilon_0 \kappa}. \tag{26.24}$$

This result—Gauss's law in dielectrics—is remarkable. The left side contains the electric flux of the electric field *inside the dielectric*. We can obtain this field, however, just by accounting for the enclosed *free* charge (and we already know how to deal with that charge). This relationship is valid because the effect of the bound charge is completely accounted for by the dielectric constant in the denominator. As Eq. 26.17 shows, dividing q_{free} by κ gives the difference of the free and bound charges.

Because the dielectric constant affects the value of the electric field, **Gauss's law in matter** is usually written in the form

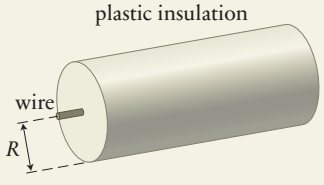
$$\oint \kappa \vec{E} \cdot d\vec{A} = \frac{q_{\text{free, enc}}}{\epsilon_0}. \tag{26.25}$$

This form of Gauss's law is very general: Even though we derived it for the special case of a parallel-plate capacitor, it holds in any situation, even one without a dielectric. In the absence of matter (that is, in vacuum), $\kappa = 1$, and because there is no bound charge we have $q_{\text{free, enc}} = q_{\text{enc}}$. Then Eq. 26.25 becomes identical to the familiar form of Gauss's law (Eq. 24.8).

EXAMPLE 26.7 Electric field surrounding a charged insulated wire

A thin, long, straight wire is surrounded by plastic insulation of radius R and dielectric constant κ (Figure 26.32). The wire carries a uniform distribution of charge with a positive linear charge density λ. If the wire has a diameter d, what is the potential difference between the outer surface of the wire and the outer surface of the insulation?

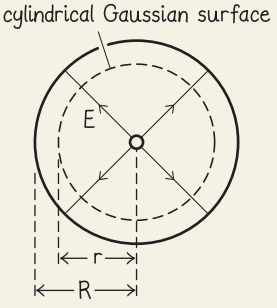
Figure 26.32 Example 26.7.



❶ **GETTING STARTED** The insulation reduces the electric field created by the charge in the wire, so the potential difference between the wire surface and any location a distance R from the wire (in other words, any location on the outer surface of the insulation) is smaller than when there is no insulation around the wire.

❷ **DEVISE PLAN** The potential difference between two locations A and B can be obtained from Eq. 25.25, $V_{AB} = -\int_A^B \vec{E} \cdot d\vec{\ell}$, but using this expression requires me to know the electric field. To calculate the electric field inside the insulation, I can apply Eq. 26.25 to a cylindrical Gaussian surface that has radius r and length L and is concentric with the wire, as shown in Figure 26.33.

Figure 26.33



QUANTITATIVE TOOLS

3 EXECUTE PLAN Because of the cylindrical symmetry of the wire, the electric field has the same magnitude E everywhere on the curved region of the Gaussian surface. The electric flux through that region of the Gaussian surface is equal to the product of the electric field at a distance r from the wire and the surface area: $\Phi = EA = E(2\pi rL)$ (Eq. 3 in Exercise 24.7). The free charge enclosed by the cylinder is λL, and with these substitutions Eq. 26.25 becomes

$$\kappa E(2\pi rL) = \frac{\lambda L}{\epsilon_0}$$

and

$$E = \frac{\lambda}{2\pi\kappa\epsilon_0 r}. \qquad (1)$$

Substituting this expression for E into Eq. 25.25, I obtain for the potential difference between the outer surface of the wire and the outer surface of the insulation

$$V_{dR} = -\int_{d/2}^{R} \vec{E}\cdot d\vec{r} = -\frac{\lambda}{2\pi\kappa\epsilon_0}\int_{d/2}^{R}\frac{1}{r}\,dr$$
$$= -\frac{\lambda}{2\pi\kappa\epsilon_0}\ln\frac{2R}{d}. \checkmark$$

4 EVALUATE RESULT Because $\ln(2R/d)$ is positive, the potential difference is negative, as it should be because I am moving away from the positively charged wire. As an additional check, if I set $\kappa = 1$ in Eq. 1, my result for the electric field becomes identical to the result I obtained for a thin wire without insulation in Exercise 24.7.



26.23 Show that, if you account for the free and bound charges, Gauss's law in vacuum (Eq. 24.8) yields the same result for the electric field outside the insulation as Gauss's law in matter (Eq. 26.25) does.

Chapter Summary

Capacitors (Sections 26.1, 26.2, 26.5)

Concepts A **charge-separating device** (such as a battery) has some mechanism that moves charge carriers *against* an electric field. The work done in this process increases the system's electric potential energy.

A **capacitor** consists of a pair of conducting objects separated by a nonconducting material or vacuum. The objects store electric potential energy once charge has been transferred from one to the other.

A *parallel-plate capacitor* consists of two parallel conducting plates of surface area A separated by a gap of width d. The electric field is uniform between the plates.

A *coaxial capacitor* consists of two coaxial conducting cylinders of radii R_1 and $R_2 > R_1$ and length $\ell \gg R_2$.

A spherical capacitor consists of two concentric conducting spherical shells of radii R_1 and $R_2 > R_1$.

Quantitative Tools The conducting objects in a capacitor carry charges of equal magnitude q but opposite sign. The charge separation produces a potential difference of magnitude V_{cap} between the objects. The **capacitance** C of this arrangement is

$$C \equiv \frac{q}{V_{\text{cap}}}. \tag{26.1}$$

Capacitance is measured in **farads** '(F),' where

$$1\ \text{F} \equiv 1\ \text{C/V}.$$

The capacitance of a *parallel-plate capacitor* is

$$C = \frac{\epsilon_0 A}{d}.$$

The capacitance of a *coaxial capacitor* is

$$C = \frac{2\pi\epsilon_0 \ell}{\ln(R_2/R_1)}.$$

The capacitance of a spherical capacitor is

$$C = 4\pi\epsilon_0\left[\frac{R_1 R_2}{R_2 - R_1}\right].$$

Electric field energy and emf (Sections 26.4, 26.6)

Concepts The **energy density** of an electric field is the energy per unit volume stored in the field.

The **emf** of any charge-separating device is the work per unit charge done by nonelectrostatic interactions in separating positive and negative charge carriers.

Quantitative Tools The electric potential energy U^E stored in a capacitor is

$$U^E = \tfrac{1}{2}\frac{q^2}{C} = \tfrac{1}{2}CV_{\text{cap}}^2 = \tfrac{1}{2}qV_{\text{cap}}. \tag{26.4}$$

In air or vacuum, the **energy density** u_E of an electric field is

$$u_E = \tfrac{1}{2}\epsilon_0 E^2. \tag{26.6}$$

The **emf** $\mathscr{E}$ of a charge-separating device is

$$\mathscr{E} \equiv \frac{W_{\text{nonelectrostatic}}}{q}. \tag{26.7}$$

Dielectrics (Sections 26.3, 26.7, 26.8)

Concepts A **dielectric** is a polarizable nonconducting material. A *polar* dielectric is made up of molecules that have a permanent dipole moment, whereas a *nonpolar* dielectric consists of molecules that do not have a dipole moment in the absence of an electric field.

Quantitative Tools The **dielectric constant** κ of a material between capacitor plates is

$$\kappa \equiv \frac{V_0}{V_{\text{d}}}, \tag{26.9}$$

where V_0 is the potential difference across the capacitor without the dielectric and V_{d} is the potential with the dielectric in place.

A dielectric inserted between the plates of a capacitor becomes polarized by the electric field of the capacitor. This polarization gives the two faces of the dielectric thin layers of charge of equal magnitude but opposite sign. This charge is **bound** because the charge carriers are not free to move. The charge on the capacitor plates is **free** because the charge carriers are free to move.

If a capacitor has capacitance C_0 without a dielectric, its capacitance with a dielectric is

$$C_d = \kappa C_0. \tag{26.11}$$

If q_{free} is the **free charge** on a capacitor, then the **bound charge** q_{bound} is

$$q_{bound} = \frac{\kappa - 1}{\kappa} q_{free}. \tag{26.18}$$

Gauss's law in matter is

$$\oint \kappa \vec{E} \cdot d\vec{A} = \frac{q_{free,\,enc}}{\epsilon_0}, \tag{26.25}$$

where $q_{free,\,enc}$ is the free charge enclosed by the Gaussian surface.

Questions and Problems

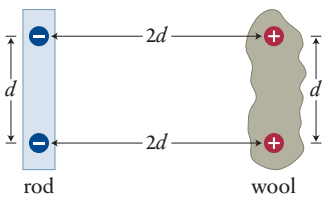
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

26.1 Charge separation

1. Is it possible separate positive and negative charges unequally in two objects by a charge-separating method or by any mechanical equipment? •
2. Positive work is done on a system containing positively and negatively charged particles. All work goes into changing the electric potential energy of the system. What can be said about the electric field between positively and negatively charged particles? What can be said about the electric field everywhere else? ••
3. If you are designing a Van de Graaff generator, would you prefer to replace the spherical metallic dome with a dome of some other shape to achieve better charge storage? Justify your answer. ••
4. Draw energy diagrams for the processes of separating charge carriers by (*a*) rubbing, (*b*) increasing the separation distance between a positively charged object and a negatively charged object, and (*c*) increasing the separation distance between two positively charged objects. ••
5. A plastic rod is rubbed with wool, producing a distribution of positive and negative surplus charge that is concentrated in two locations on the rod and two on the wool. This distribution can be approximated as two pairs of small balls of charge as shown in Figure P26.5. The distance d is 30.0 mm, and the magnitude of the charge on each ball is 2.00 μC. If 34.0 J of work went into thermal energy as the rod and wool heated up, how much work was done on the system? ••

Figure P26.5



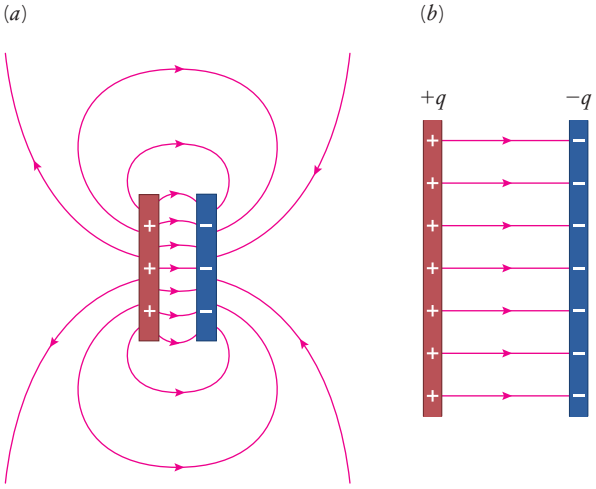
6. A piece of wool is used to charge two plastic spheres of negligible radius. When the spheres are held 20 mm apart, they repel each other with 16 N of force. If the wool ends up with a surplus positive charge of 30 μC, what is the charge on each of the two spheres? ••
7. Two 0.0450-kg spheres are identical (including the charge they carry) and are initially pinned in place 200 mm apart. You unpin them and push on each sphere with a force of 0.15 N. After their separation increases to 400 mm, each sphere has a speed of 500 mm/s. What is the charge on each sphere? Is there more than one possible answer? ••
8. A fellow scientist heard that a Van de Graaff generator built 70 years ago could collect 5.0 C of charge on its dome, which had a radius of 1.1 m, and has challenged you to do the same. The belt you plan to use is 100 mm wide and 10.0 m long (5.0 m to go up to the dome, and 5.0 m to come back down). Charging the belt gives it a surface charge density of 45 μC/m^2. How much force must your motor be able to exert on the belt in order to accomplish your goal? If you had more freedom in the design, what changes would you make to make the charging process easier on your motor, and why? Would your changes alter only the force required for your motor, or would they also affect the energy required for charging? State any assumptions you make, and state whether or not they are physically reasonable. •••

26.2 Capacitors

9. The plates of a capacitor are charged using a battery, and they produce an electric field across the separation distance d between them. The two plates are now to be pushed together to a separation of $d/2$. The pushing together can be done either with the battery connected or with it disconnected. Which way would result in the greater electric field magnitude, and by what factor? •
10. Two parallel-plate capacitors have the same plate area. Capacitor 1 has a plate separation twice that of capacitor 2, and the quantity of charge you place on capacitor 1 is twice the quantity you place on capacitor 2. How do the potential differences across each of the two capacitors compare to each other? •

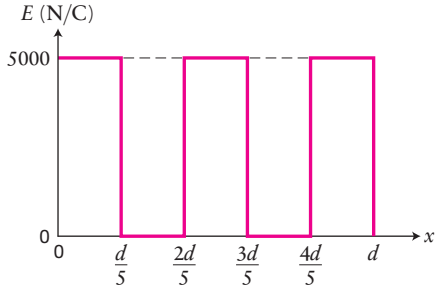
11. Two parallel-plate capacitors have the same plate area. Capacitor 1 has a plate separation twice that of capacitor 2, and the potential difference you impose across the plates of capacitor 1 is twice the potential difference you impose across capacitor 2. How do the quantities of charge stored on the two capacitors compare to each other? •
12. Suppose a capacitor is fully charged by a battery and then disconnected from the battery. The positive plate has a charge $+q$ and the negative plate has a charge $-q$. The plate area is doubled, and the plate separation is reduced to half its initial separation. What is the new charge on the negative plate? ••
13. Explain why there must always be some nonzero electric field outside the plates of a parallel-plate capacitor (Figure P26.13*a*), which means that the idealization shown in Figure P26.13*b*, with all the field lines confined to the region between the plates, can never be exact. (Hint: Consider the potential difference between the plates for paths outside the edges.) ••

Figure P26.13



14. Figure P26.14 is a graph of the electric field between two capacitor plates as a function of distance from the left plate. The distance between the plates is d, the direction of the field is to the right, and a battery is connected to the plates. (*a*) Which plate carries a positive charge? (*b*) How many slabs of material have been inserted between the plates? (*c*) What can you say about the slab(s) inserted? (*d*) Make a graph of the electric field between the plates when the slabs between the plates are removed and the battery remains connected. ••

Figure P26.14



15. When two parallel plates separated by a distance d are connected to a battery that maintains a potential difference of magnitude V between the plates, a charge of magnitude q accumulates on each plate. If the plate separation is increased to $2d$ and the battery is replaced with one that has potential difference $2V$ between the terminals, what is the new charge on each plate? ••
16. Suppose two capacitor plates have an area of 1 cm^2 and are initially separated by 2 mm. Each plate holds 5 μC of charge. How much energy is required to double the plate separation? ••
17. A certain capacitor is fully charged by a battery, such that the positive plate holds a charge $+q$ and the negative plate holds a charge $-q$. The plates of the capacitor are then pulled apart to twice their initial separation. Determine the new charge on the positive plate if (*a*) the battery is kept attached to the capacitor as the separation is increased, and (*b*) the battery is disconnected before the plate separation is increased. ••
18. You need to insert a metal slab between the two plates of a parallel-plate capacitor. The plates are a distance d apart, and a battery maintains a constant potential difference V_{batt} between them. In order to avoid dielectric breakdown, the electric field in any region cannot exceed $4V_{batt}/d$. (*a*) What is the maximum thickness of the metal slab that can be used without causing dielectric breakdown? (*b*) If this maximum thickness is used, does the insertion of the metal slab affect the amount of charge stored on the capacitor? •••

26.3 Dielectrics

19. Either a dielectric or a conductor could be inserted between the plates of a capacitor. (*a*) State at least two similarities between the effects of inserting a dielectric and inserting a conducting slab. (*b*) State at least two differences between the effects of inserting a dielectric and inserting a conducting slab. •
20. A fully charged capacitor initially has an air gap and is disconnected from the battery. A dielectric material is inserted between the plates. What happens to (*a*) the free charge at the surface of the capacitor plates and (*b*) the total charge (free and bound) at the surface of the capacitor plates? •
21. Two parallel-plate capacitors have the same dimensions, but the space between the plates is filled with air in capacitor 1 and with plastic in capacitor 2. The magnitude of the charge on the plates is the same in both capacitors. Compare (*a*) the magnitudes of the electric fields E_1 and E_2 between the plates, (*b*) the potential differences V_1 and V_2 between the plates, and (*c*) the energies stored in the capacitors. ••
22. Two parallel-plate capacitors have the same dimensions, but the space between the plates is filled with air in capacitor 1 and with plastic in capacitor 2. The potential difference between the plates is the same in both capacitors. Compare (*a*) the magnitudes of the electric fields E_1 and E_2 between the plates, (*b*) the magnitudes of the charges on the plates q_1 and q_2, and (*c*) the energies stored in the capacitors. ••

23. A capacitor connected to a battery initially holds a charge of $+q$ on its positive plate and $-q$ on its negative plate. The electric field between the plates is initially $\vec{E}$. A dielectric material is then inserted that polarizes in such a way as to produce an electric field of $-0.50\vec{E}$ (where the – sign indicates that this field opposes the initial field $\vec{E}$). Determine the new charge stored on the positive plate. ••
24. A parallel-plate capacitor with air between the plates and plate separation distance d is connected to a battery that maintains a potential difference V_{batt} between the plates. Sketch a graph of V as a function of x when x is defined as position between the plates, running from $x = 0$ at the surface of the positively charged plate to $x = d$ at the surface of the other plate. Plot the points $x = 0, x = d/4, x = d/2, x = 3d/4$, and $x = d$. A slab that is made of a dielectric material and has a thickness of $d/2$ is inserted between the plates such that the midpoint of the slab thickness is at the position $d/2$ in the space between the plates. Sketch a graph of $V(x)$ for the capacitor now. Comment on how your graph is affected by the magnitude of the bound charge induced on the surface of the dielectric slab. ••
25. A capacitor initially has a charge of magnitude q on each plate. When a dielectric is inserted between the plates, the bound surface charge on the two dielectric surfaces facing the plates has a magnitude $q/4$. What is the ratio of the electric field magnitude in the empty capacitor to the electric field magnitude in the capacitor with the dielectric (*a*) if the battery stays connected to the plates and (*b*) if the battery is removed before the dielectric is inserted? ••
26. A capacitor has a plate area of 0.0045 m^2 and a charge of magnitude q on each plate. The space between plates has been filled with a dielectric that has a bound surface charge of magnitude $3q/4$ on either side. When the dielectric material is removed, the electrical breakdown of air is observed between the plates. What is the smallest possible value of q? ••
27. The night before an exam, your study partner asks what you can do to a parallel-plate capacitor to avoid the problem of the electrical breakdown of air. You answer that a nonconducting dielectric material can be inserted. Your friend is confused as to why other things wouldn't work. Why wouldn't it help, for example, to add a conducting slab? Why isn't increasing plate separation also a valid answer? After all, moving charge carriers very far apart must decrease their effect on the air between them. Why do these alternatives not work? •••

26.4 Voltaic cells and batteries

28. The term *emf* (electromotive force) represents work done per unit charge. Why, then, is it termed erroneously as "force" (or electromotive force)? •
29. A 10-V battery is used for a capacitor to get charged with air inside the plates. If the air is replaced with mica, will the voltage of the battery reduce if it is given that mica has higher dielectric constant value than air? •
30. What limits the lifetime of a battery? •
31. The reactions in a lead-acid battery are

$$\text{Positive terminal:}\quad PbO_2 + HSO_4^- + 3H^+ + 2e^- \rightarrow PbSO_4 + 2H_2O$$

$$\text{Negative terminal:}\quad HSO_4^- + Pb \rightarrow PbSO_4 + H^+ + 2e^-$$

If a certain battery supplies 1.0×10^8 electrons per second to the negative terminal and the battery contains 3.00 moles of electrolytic solution (which means the solution contains 3.00 moles of HSO_4^-), what fraction of the solution undergoes a chemical reaction each second? ••
32. One electrode in an alkaline battery is zinc, and the other is manganese dioxide. Fill in the blank for the chemical reaction (*a*) at the zinc electrode: $Zn + 2OH^- \rightarrow ZnO + ____ + 2e^-$ and (*b*) at the manganese electrode: $2MnO_2 + H_2O + ____ \rightarrow Mn_2O_3 + 2OH^-$. (*c*) Which electrode is positive and which is negative? ••
33. Suppose that the metal electrodes in a battery are shaped like large parallel plates and are 2.00 mm apart so that they are essentially a parallel-plate capacitor. If the battery supplies a potential difference of 1.5 V between the electrodes, what is the surface charge density on the electrodes (before anything has been connected to the battery)? ••

26.5 Capacitance

34. A capacitor having capacitance 10 pF is connected to a 9-V battery. If you consider a cylindrical Gaussian surface around the capacitor, calculate the net electric flux through it when the capacitor is operational. •
35. Can you increase the capacitance of a parallel plate capacitor to infinity? Explain your answer. •
36. A parallel-plate capacitor initially has capacitance C. What is the capacitance after each of the following single changes is made: (*a*) plate separation distance doubled, (*b*) dielectric material inserted, (*c*) potential difference across plates doubled, (*d*) plate area halved, (*e*) applied emf increased until plate charge doubles. •
37. A coaxial capacitor consisting of an inner wire and an outer cylindrical shell has a length of 50 cm and an outer diameter of 4 mm. When connected to a 16.0-V battery, this capacitor can hold 3 nC of charge on the wire. What is the wire radius? ••
38. How is Farad is a high unit of measuring the capacitance? Explain your argument by considering a parallel-plate capacitor having a plate separation of (*d*) 20 cm and plate area Am^2. ••
39. One way to double the capacitance of a parallel-plate capacitor is to reduce the plate separation distance by a factor of 2. Doubling the capacitance of a spherical capacitor is a bit more complex, though. (*a*) If you had a spherical capacitor consisting of two concentric spherical shells of radii R_{inner} and $R_{outer} > R_{inner}$, you could

double the capacitance by changing the radius of the outer shell from R_{outer} to some value d. What is d in terms of R_{inner} and R_{outer}? (*b*) You could also double the capacitance by changing the radius of the inner shell from R_{inner} to some value d. What is d in terms of R_{inner} and R_{outer}? (*c*) Repeat parts *a* and *b* for a capacitor consisting of two long concentric cylindrical shells of radii R_{inner} and $R_{outer} > R_{inner}$. ••

40. A spherical capacitor consists of two concentric conducting spherical shells of radii R and $2R$. (*a*) How long would a coaxial cylindrical capacitor made of two concentric cylindrical conductors of radii R and $2R$ have to be in order to have the same capacitance as the spherical capacitor? (*b*) What plate separation distance would be required for a parallel-plate capacitor made of two circular conductors each of radius R in order to have the same capacitance as the spherical capacitor? ••

41. Two small, irregularly shaped conducting objects, one carrying charge $+q$ and one carrying charge $-q$, are placed on an x axis at $x = -4.0$ m and $x = +4.0$ m, respectively. Between the objects, the electric field on the x axis is given by $E(x) = aq(x^2 + b)$, where a and b are constants. (*a*) What are the standard units of a and b? (*b*) What is the capacitance of this configuration? •••

42. The expression for the capacitance of a parallel-plate capacitor that has plates of area A separated by a distance d is given in Example 26.2. (*a*) Show that the expression for a spherical capacitor, given in Example 26.4, reduces to the expression for a parallel-plate capacitor in the limit that R_1 approaches R_2, with $d = R_2 - R_1$. (*b*) Show that the expression for a cylindrical capacitor, given in Example 26.3, reduces to the expression for a parallel-plate capacitor in the limit that R_1 approaches R_2, with $d = R_2 - R_1$. (Hint: You might find the Taylor series expansion of $\ln(1 + x)$ useful.) •••

26.6 Electric field energy and emf

43. In a parallel plate capacitor, transfer of 2×10^{36} electrons takes place when connected to a 12-V battery. Estimate the energy stored in the capacitor. •

44. Suppose a certain battery has an internal emf of 15 V but the potential difference across its terminals is only 80% of that value. If that battery is connected to a 47 μF capacitor, how much energy is stored when the capacitor is fully charged? •

45. A parallel-plate capacitor connected to a battery maintaining a potential difference V across the capacitor initially stores electric potential energy U_i^E. If the plate area is doubled and the battery is replaced with one that maintains a potential difference of $2V$ across the capacitor, how much electric potential energy is stored now? •

46. Two parallel-plate capacitors 1 and 2 are identical except that capacitor 1 has charge $+q$ on one plate and $-q$ on the other, and capacitor 2 has charge $+2q$ on one plate and $-2q$ on the other. Compare the two capacitors based on their (*a*) capacitance, (*b*) potential difference between the plates, (*c*) electric field magnitude between the plates, and (*d*) energy stored. ••

47. Three parallel-plate capacitors are separately connected to identical batteries. Capacitor 1 has a plate area A and a plate separation d. Capacitor 2 has a plate area $2A$ and a plate separation d. Capacitor 3 has a plate area A and a plate separation $2d$. Rank the three capacitors, largest first, based on (*a*) capacitance, (*b*) charge stored, (*c*) electric field magnitude between the plates, (*d*) energy stored, and (*e*) energy density. ••

48. Three parallel-plate capacitors each store the same amount of charge. Capacitor 1 has a plate area A and a plate separation d. Capacitor 2 has a plate area $2A$ and a plate separation d. Capacitor 3 has a plate area A and a plate separation $2d$. Rank the three capacitors, largest first, based on their (*a*) capacitance, (*b*) potential difference between the plates, (*c*) electric field magnitude between the plates, (*d*) energy stored, and (*e*) energy density. ••

49. (*a*) A spherical capacitor consists of two concentric conducting spheres of radii R and $2R$. If the two spheres carry charges of $+q$ and $-q$, what is the average energy density inside the capacitor in terms of R and q? (*b*) A cylindrical capacitor of length ℓ consists of two coaxial conducting cylinders of radii R and $2R$. If the two cylinders carry charges of $+q$ and $-q$, what is the average energy density inside the capacitor? ••

50. A parallel-plate capacitor has plates of area A. The plates are initially separated by a distance d, but this distance can be varied. If the capacitor is connected to a battery, what should the plate separation be if you want to quadruple (*a*) the capacitance, (*b*) the charge stored, (*c*) the energy stored, and (*d*) the energy density? ••

51. A parallel-plate capacitor has plates of area A. The plates are initially separated by a distance d, but this distance can be varied. If the capacitor is charged by a battery and the battery is then removed so that the capacitor is isolated, what should the plate separation be if you want to quadruple (*a*) the capacitance, (*b*) the charge stored, (*c*) the energy stored, and (*d*) the energy density? ••

52. A parallel-plate capacitor has an initial charge q and a plate separation distance d. How much work must you do, in terms of q, d, and plate area A, to increase the separation distance to $3d$ if (*a*) the capacitor is isolated and (*b*) the capacitor is connected to a battery? (*c*) The work you do is positive in both cases. Explain why in one situation the electric potential energy increases and in the other it decreases. •••

26.7 Dielectric constant

53. If you are constructing a parallel-plate capacitor, having air and paper as the dielectric medium between the plates, which dielectric material within the plates will make the capacitor act better? Justify your answer. •

54. The potential difference across the plates of a parallel-plate capacitor is gradually increased. (*a*) What happens when the breakdown threshold is exceeded? (*b*) If you want to increase the potential difference before breakdown occurs, what capacitor variables might you change? •

55. (*a*) Make a list of ways to increase the capacitance of a parallel-plate capacitor. (*b*) Does *Increase quantity of charge stored* or *Decrease potential difference* belong on the list? •

56. When you remove a dielectric slab from between the plates of a charged isolated capacitor, what happens to the energy stored in the capacitor? Why does this happen to the stored energy? ••

57. Two parallel-plate capacitors are identical except that capacitor 1 has vacuum between the plates and capacitor 2 has a dielectric slab of dielectric constant κ filling the space between the plates. The capacitors are connected to identical batteries. Compare the two based on (*a*) capacitance, (*b*) charge stored, (*c*) energy stored, (*d*) electric field magnitude, and (*e*) energy density. ••

58. Two parallel-plate capacitors are identical except that capacitor 1 has vacuum between the plates and capacitor 2 has a dielectric slab of dielectric constant κ filling the space between the plates. Each capacitor is isolated (that is, not connected to a battery), and they store equal quantities of charge. Compare the two based on (*a*) capacitance, (*b*) potential difference between the plates, (*c*) energy stored, (*d*) electric field magnitude between the plates, and (*e*) energy density. ••

59. When a dielectric slab completely fills the space between the plates of a parallel-plate capacitor, the magnitude of the bound charge is one-fourth the magnitude of the free charge. (*a*) What is the dielectric constant of the material of which the slab is made? (*b*) If the slab does not fill the space between the plates, but instead has a width equal to half the distance between the plates, does the magnitude of the bound charge change? If so, how? ••

60. If you want to maintain a potential difference of 6000 V between the plates of a parallel-plate capacitor, what is the minimum value of the plate separation if the space between the plates is completely filled by (*a*) a slab of mica and (*b*) a slab of barium titanate? (*c*) What is the ratio of the maximum charges that could be put on a plate of the capacitor in part *b* to the capacitor in part *a*. ••

61. A parallel-plate capacitor with air between its plates can hold 6 pC of charge per volt of potential difference across its plates. When a barium titanate dielectric slab completely fills the space between the plates and the capacitor is connected to a 9.0-V battery, what is the charge that can be stored on the positive plate of the capacitor? ••

62. You are working on charge-storage devices for a research center. Your goal is to store as much charge on a given device as possible. The facilities allow you to generate almost any potential difference V you need, but you are restricted to using a single parallel-plate capacitor. The area and separation distance of the plates are fixed, and the dielectric materials available to you are paper ($\kappa = 3.0$, $E_{max} = 4.0 \times 10^7$ V/m), Mylar ($\kappa = 3.3$, $E_{max} = 4.3 \times 10^8$ V/m), quartz ($\kappa = 4.3$, $E_{max} = 8 \times 10^6$ V/m), and mica ($\kappa = 5$, $E_{max} = 2 \times 10^8$ V/m). What properties of these materials must you consider in choosing the best dielectric for your needs? Rank the materials in order of their ability to meet your needs, first choice first. ••

63. A spherical capacitor has an inner radius of 5 cm and an outer radius of 6 cm. With air between the spheres, the capacitor is connected to a 12-V battery and allowed to charge fully. With the battery still connected, oil ($\kappa = 2.2$ is poured in, filling the volume between the spheres. Calculate the additional work done by the battery throughout the process of filling the capacitor with oil. •••

64. A dielectric slab completely fills the space between the plates of a parallel-plate capacitor. The magnitude of the bound charge on each side of the slab is 75% of the magnitude of the free charge on each plate. The capacitance is $480\epsilon_0\ell$, where ℓ is a constant with dimensions of length, and the maximum charge that can be stored on the capacitor is $240\ell^2\epsilon_0 E_{max}$, where E_{max} is the breakdown threshold. What are (*a*) the dielectric constant for the slab, (*b*) the plate separation distance in terms of ℓ, and (*c*) the plate area in terms of ℓ? •••

26.8 Gauss's law in dielectrics

65. Two very long wires each carry a linear charge density λ. They initially repel each other with a force F. If the wires are immersed in distilled water, with what force do they repel each other? •

66. A solid conducting sphere of radius R and carrying charge $+q$ is embedded in an electrically neutral nonconducting spherical shell of inner radius R and outer radius $2R$. The material of which the shell is made has a dielectric constant of 4.0. (*a*) Relative to a potential of zero at infinity, what is the potential at the center of the conducting sphere? (*b*) If the dielectric constant of the shell doubles, what happens to the potential at the center of the sphere? ••

67. A solid conducting sphere of radius $R = 10$ cm is embedded in an electrically neutral nonconducting spherical shell that has inner radius R, outer radius $2R$, and is made of a material having dielectric constant κ. The potential at the center of the sphere is 540 V. The electric field at the outer surface of the shell is directed away from the surface of the shell, and has a magnitude of 900 V/m. What are (*a*) the dielectric constant of the shell and (*b*) the electric charge on the sphere? ••

68. A conducting sphere has a radius of 2.25 m and carries a positive surplus charge of 35.0 mC. A protective layer of barium titanate is applied to the surface of the sphere to make it safe for laboratory workers nearby. Safety considerations dictate that the potential difference between the surface of the conductor and the outside of the nonconductive layer must be 20,000 V (though this is by no means safe under all conditions). How thick must the protective layer be? •••

Additional Problems

69. (*a*) Explain how treating the electric field lines between two oppositely charged objects as elastic bands can help you to understand what happens to the electric potential energy of the system when you increase the separation distance between the objects. (*b*) Can you think of any limitations to this elastic-band analogy for analyzing field lines? •

70. Why are capacitors used in external defibrillators instead of high-voltage batteries? •

71. Two parallel-plate capacitors have the same plate separation. The plate area of capacitor 1 is twice that of capacitor 2. (*a*) How do the potential differences across the two capacitors compare if the quantity of charge you place on capacitor 1 (q_1) is twice the quantity you place on capacitor (q_2)? (*b*) How do the charge magnitudes on either plate of the two capacitors compare if the potential difference between the plates of capacitor 1 is twice that between the plates of capacitor 2? •

72. During a lightning strike, on the order of 10 C of charge is typically transferred to the ground over a potential difference of 3×10^8 V. (*a*) What is the capacitance of the cloud-ground system? (*b*) How much energy is stored in the system just before the strike? (*c*) As a check on the order of magnitude of your answer for part *b*, convert the energy released in the lightning strike into liters of gasoline (gasoline has a stored chemical energy of 36 MJ/liter). ••

73. Given that charge separation increases the electric potential energy of a system, what can you conclude about the criteria for keeping a system of positively and negatively charged particles in static equilibrium, in the presence of electrostatic interactions only? ••

74. A parallel-plate capacitor in which the plates are extendable is connected to a battery and charged until there is a charge $+q$ on one plate and a charge $-q$ on the other plate. (*a*) With the battery still connected, the area of each plate is doubled by sliding the extension out. How much charge is on each plate now? (*b*) In which case, before the plate area is expanded or after, is more force required to keep the plates apart? Use the elastic-band model described in Problem 69 to explain your answer. ••

75. A parallel-plate capacitor carries a charge $+q$ on one plate and a charge $-q$ on the other plate. Each plate has an area A. How much force, in terms of q and A, does one plate exert on the other? ••

76. When a parallel-plate capacitor has air between its plates and is connected to a 12.0-V battery, its capacitance is 33 n*F*. The capacitor is then submerged in distilled water. What is the charge on the plates (*a*) if the capacitor is submerged with the battery connected and (*b*) if the battery is removed before submersion? ••

77. Your firm uses a large parallel-plate capacitor to store energy, and you measure the electric field strength between the plates to determine the amount of energy stored. During a test run with a new renewable energy source, you realize that the electric field magnitudes being produced are too small to be accurately detected by your equipment. The capacitor plates are fixed in place, and you cannot yet control the energy source to give you a greater potential difference. All you have to work with are four large flat slabs of metal, one $d/2$ thick, one $d/3$ thick, one $d/4$ thick, and one $d/5$ thick, where d is the fixed plate separation distance. You realize that it is necessary to increase E by at least an order of magnitude, but it takes you a while to figure out what to do. How do you increase E sufficiently? ••• CR

78. You are working for a car-battery manufacturer, and your boss complains that to be more competitive, the company needs to produce batteries that will last longer while minimizing the amount of metal or metal oxide used. She is not concerned about the amount of electrolyte used, but the battery must be powerful enough to supply a great deal of energy quickly when needed. What do you think the shape of the electrodes might have to do with battery life and power, starting with what seem to be the extremes: electrodes shaped like sheets similar to parallel plates and rod-shaped electrodes that extend down into the electrolyte solution? ••• CR

79. In the endless endeavor to make electronic devices as small as possible, you have been hired to make a capacitor that has the greatest capacitance possible in a cubic volume of $(10\text{ mm}) \times (10\text{ mm}) \times (10\text{ mm})$. You are allowed to use any capacitor geometry you wish, as well as any combination of geometries. For simplicity, assume a dielectric constant of 1000 (as better materials are found, you can multiply the capacitance appropriately). You begin thinking about the minimum thickness of the conductors and dielectrics, and the scale of about 0.5 μm seems reasonable. There is the question of how to fill the volume while maintaining exactly two conducting surfaces. Describe the capacitor that meets all of these criteria. ••• CR



27

Magnetic Interactions

27.1 Magnetism
27.2 Magnetic fields
27.3 Charge flow and magnetism
27.4 Magnetism and relativity

27.5 Current and magnetism
27.6 Magnetic flux
27.7 Moving particles in electric and magnetic fields
27.8 Magnetism and electricity unified

CONCEPTS

QUANTITATIVE TOOLS

The word *magnetism* is derived from Magnesia, a province where the ancient Greeks mined *magnetite*, a mineral that attracts iron. The interactions between a magnet and a paper clip or a refrigerator door are familiar ones. These interactions may seem to have nothing to do with the subject of earlier chapters, but as we shall see, electricity and magnetism are closely related phenomena. They are two manifestations of one interaction, called the *electromagnetic interaction*. The discovery of the connection between electricity and magnetism in the 19th century opened the door to important technological breakthroughs, such as electric motors and generators, the transmission of radio signals, and the electronics and telecommunications industries.

In this chapter, we discuss interactions between magnets and introduce the concept of a magnetic field. We also begin to explore the connection between electricity and magnetism.

27.1 Magnetism

One simple definition of **magnet** is any object that attracts pieces of iron, such as iron filings or paper clips. Magnets come in many shapes and sizes, as Figure 27.1 shows.

If you examine the surface of a magnet interacting with an object that contains iron, such as a paper clip, you will discover that certain parts of the magnets, called *magnetic poles*, interact more strongly with the paper clip than do other parts of the magnet. Disk-shaped magnets (Figure 27.1*b*) usually have poles on the two faces; paper clips stick to the flat faces but not to the curved surface. Bar magnets (Figure 27.1*c*) have poles at the ends, and most of the length of the bar doesn't interact very strongly with a paper clip. A horseshoe magnet (Figure 27.1*d*) is simply a bent bar magnet with poles at the ends of the bent bar.

Figure 27.1 Magnets come in many shapes and sizes.

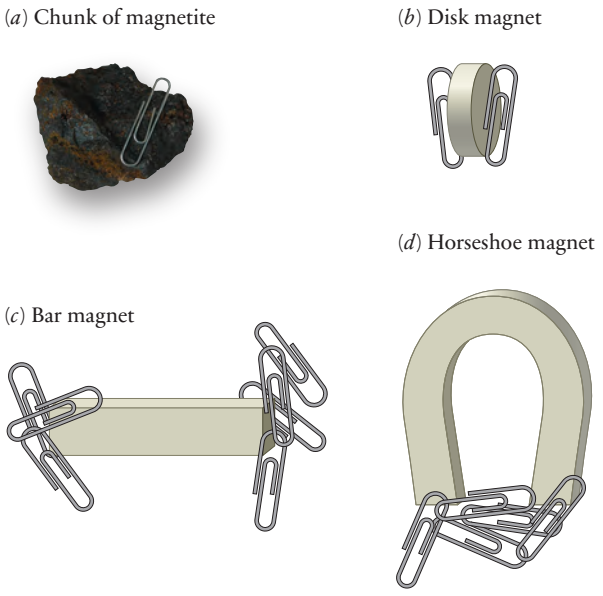


Figure 27.2 The needle of a compass is a small bar magnet that is free to rotate. When the compass is held horizontally, the needle points in the direction of Earth's North Pole. This end of the needle is defined as its north pole N.



Depending on which of their poles you hold near each another, two magnets can attract or repel without touching. This tells us that these interactions are long-range and that there are two types of magnetic poles. The interaction between magnets is one example of what are called **magnetic interactions.**

When a bar magnet is suspended and free to rotate, it aligns itself so that its poles lie on a line roughly north-south along Earth's surface. This north-south alignment of a freely rotating magnet is the basic operational principle of a compass needle, which is simply a freely rotating magnetic needle (Figure 27.2). This alignment provides a means of distinguishing between the two types of poles:

> **The pole of a freely suspended bar magnet that settles toward north is defined as being the north pole of the magnet (denoted N); the opposite pole is defined as being the south pole of the magnet (denoted S).**

To understand what causes this alignment, we must first examine how one magnetic pole interacts with another. If you hold the poles of two magnets near each other, you discover that the magnetic interaction between the two types of poles follows a rule very similar to that between the two types of charge (Figure 27.3):

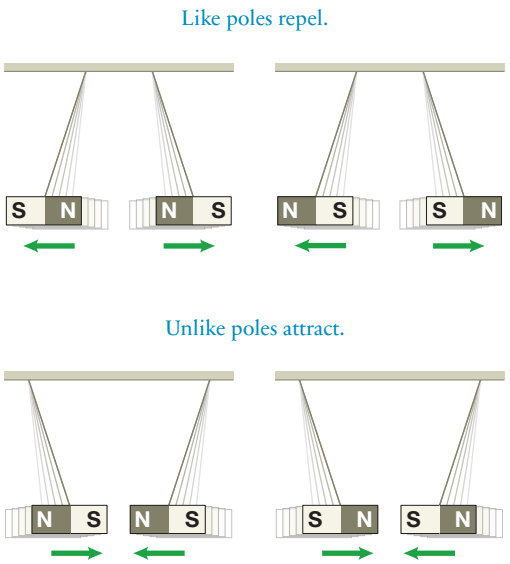
> **Like magnetic poles repel each other; opposite magnetic poles attract each other.**

CONCEPTS

27.1 Because we cannot see any obvious difference between the ends of a bar magnet, could it be that *like* poles attract each other and *unlike* poles repel each other?

How do the poles of a magnet interact with objects that are not magnets? For most materials that are not magnets, the answer to this question is: not much if at all. Try picking up a penny, a wooden tooth pick, a piece of

Figure 27.3 Interactions between magnetic poles.



aluminum, or a piece of plastic with a magnet. It won't work. However, both the north pole and the south pole of a magnet attract an iron paper clip (**Figure 27.4**).

Materials attracted by both types of magnetic poles include iron, nickel, cobalt, and certain alloys, such as steel. These materials are called *magnetic materials*. Normally, two objects made from these materials do not interact with each other. For example, one iron paper clip does not attract another one. However, just as a charged object can induce charge polarization on a neutral object, the presence of a magnet induces a **magnetic polarization** in a paper clip or any other object made from a magnetic material (**Figure 27.5**). The poles of the magnet interact with the induced poles of the now-magnetized paper clip.

27.2 (*a*) Is the interaction between a charged object and an electrically neutral object always attractive? Why or why not? (*b*) In Figure 27.4, which type of magnetic pole is induced at the top of each paper clip?

Figure 27.4 Both the north and south poles of a magnet attract an unmagnetized iron object.

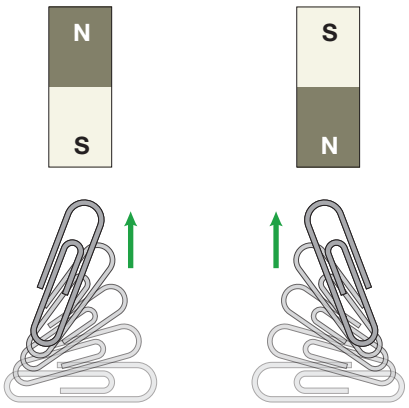
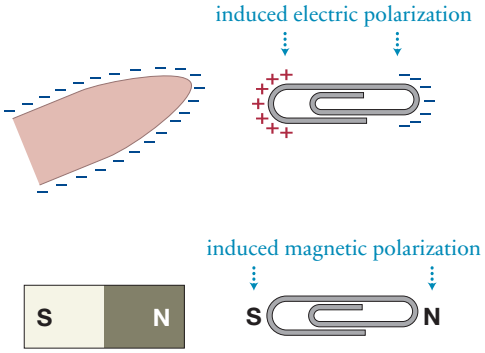


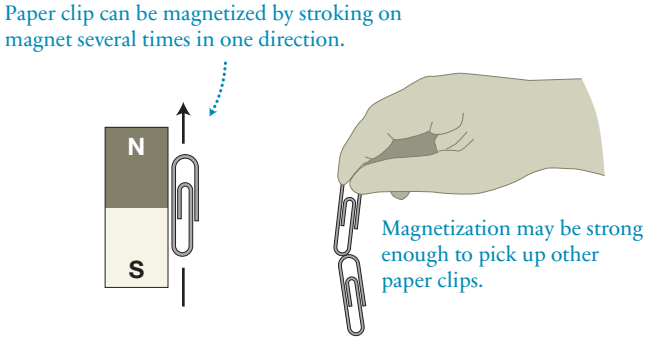
Figure 27.5 Comparing induced electric and magnetic polarization. Magnetic polarization can be induced only in an object made from magnetic material.



Unlike induced electric polarization, which vanishes as soon as the charged object is removed, some magnetic materials retain their induced magnetic polarization. For example, sewing needles left unmoved for long periods become magnetized by Earth's magnetic field. If you stroke a paper clip several times in the same direction with a magnet, the paper clip remains magnetically polarized even after you remove the magnet. In fact, you may even be able to pick up one paper clip with another (**Figure 27.6**). The clip can be *demagnetized* again—its magnetic polarization undone—by heating it, dropping it, or stroking it with a magnet in random directions.

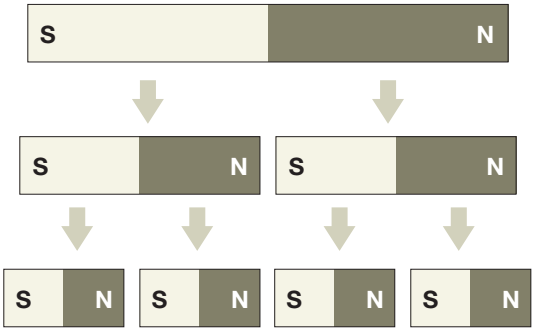
In addition to the retention of magnetic polarization, there is another fundamental difference between magnetic and electric interactions: It is not possible to isolate one pole of a magnet the way we can separate a positively charged particle and a negatively charged particle from each other. For example, if we cut a bar magnet in two, we see that each of the resulting pieces has two opposite poles (**Figure 27.7** on the next page). The cutting has created an additional *pair* of opposite poles. Remarkably, if we carefully place the two pieces together again, the two new poles seem to vanish. A piece of iron held close to the cut is either not at all or only weakly attracted, as if the two newly formed opposite poles at the cut have "neutralized" each other.

Figure 27.6 If you stroke a paper clip several times in the same direction with a magnet, the clip retains some magnetic polarization.



CONCEPTS

Figure 27.7 When a magnet is cut in two, each piece retains both an N and S pole.



Each of the two halves can again be cut in half, but regardless of how small the pieces into which we divide the bar, each piece always has two poles. As we continue to divide the bar into smaller and smaller pieces, we come to a point where each part is a single atom. Experiments show that the iron, nickel, or cobalt atoms that make up magnets still behave like magnets. Even elementary particles behave like magnets—we shall refer to these as *elementary magnets*. Thus, a magnet consists of a large number of elementary magnets whose alignment is such that their combined effect is reinforced at the poles (**Figure 27.8***a*). Because we cannot cut the elementary magnets, a pair of new poles appears every time we cut through the magnet (Figure 27.8*b*).

In spite of extensive searches, isolated magnetic poles, called *magnetic monopoles* (say, a north pole in the absence of a south pole), have never been found. Magnetism is therefore not due to magnetic monopoles, the way electricity is due to electrical charge. Although the pole of a magnet is not a monopole, we will occasionally rely on the picture of a bar magnet as a pair of opposite monopoles separated by a small distance to get an intuitive feel for magnetism. We call such an arrangement a **magnetic dipole.**

In an object made of magnetic material, the elementary magnets are randomly oriented relative to one another (**Figure 27.9***a*). Some push or pull in one direction, while others push or pull in other directions. As a result, their effects cancel and the object as a whole does not act like a magnet. When a magnet is brought nearby, however, the elementary magnets in the object align themselves with the poles of the magnet, and the object becomes magnetized (Figure 27.9*b*). With some magnetic materials, the alignment is (partially) maintained even after the nearby magnet is removed, leaving the object magnetized.

You can use this model to visualize demagnetization. If a magnet is heated or handled roughly, the elementary magnets are jarred around and lose their alignment.

Figure 27.8 The concept of elementary magnets explains why cutting a magnet in two reveals new N and S poles.

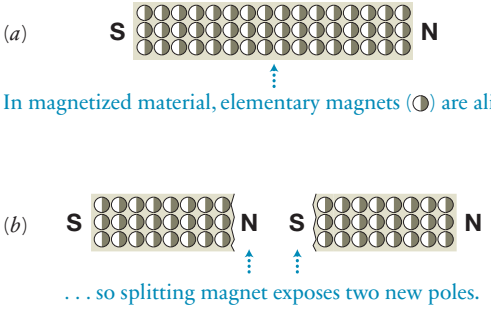
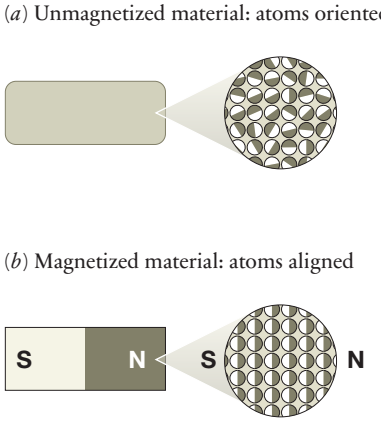
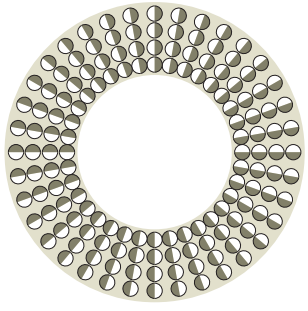


Figure 27.9 (*a*) Unmagnetized and (*b*) magnetized pieces of magnetic material.



27.3 (*a*) Draw the elementary magnets inside a bar magnet and a horseshoe magnet, using the half-filled-circle format shown in Figures 27.8 and 27.9. (*b*) How many poles does the magnetized ring in **Figure 27.10** have? (*c*) If someone gave you such a ring, how could you verify that it is indeed magnetized as illustrated?

Figure 27.10 A magnetized ring (Checkpoint 27.3).



27.2 Magnetic fields

The long-range nature of magnetic interactions suggests that we can introduce the concept of a **magnetic field,** denoted by $\vec{B}$.* In analogy to the field model for electric interactions, a magnet is surrounded by a magnetic field.

*The unintuitive symbol $\vec{B}$ was introduced early on in the description of magnetism. Whenever you see a B in an illustration or force superscript, remember that it stands for a magnetic field or force. The force exerted by a magnet on a paper clip, for example, is denoted by $\vec{F}^{B}_{mc}$.

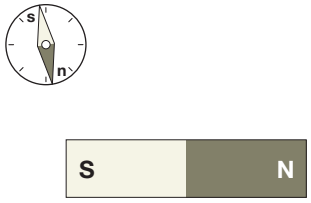
CONCEPTS

Figure 27.11 Effect of a bar magnet on a compass needle.

(*a*) Forces exerted by magnet on poles of compass needle cause a torque.



(*b*) North pole of compass needle points toward south pole of magnet



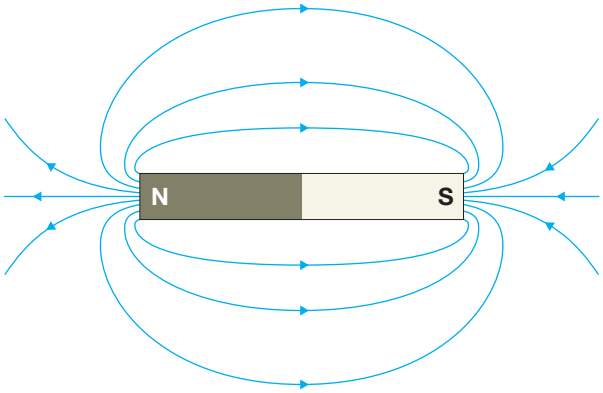
This magnetic field exerts a force on the poles of another magnet. Because there is no such thing as a "magnetic charge," it is not possible to map out a magnetic field using a "test magnetic charge." We can, however, use a compass to map out the direction of the magnetic field.

27.4 (*a*) Which end of a compass needle is a north pole: the end that points toward Earth's North Pole or the other end? (*b*) If you place a compass near the north pole of a magnet, what happens to the compass needle? (*c*) Is Earth's geographic North Pole a magnetic north pole?

When a compass is placed near a bar magnet as illustrated in **Figure 27.11a**, the south pole of the bar magnet exerts an attractive force on the north pole of the compass needle and a repulsive force on the needle's south pole. These two forces cause a torque that tends to align the needle with its north pole pointing toward the bar magnet's south pole.

27.5 (*a*) What is the effect of the north pole of the bar magnet on the compass needle in Figure 27.11*a*? (*b*) What is the combined effect of the bar magnet's north and south poles on the needle? (*c*) A compass placed in the position shown in Figure 27.11*b* aligns itself in the direction indicated. What is the direction of the vector sum of the forces exerted by the bar magnet's north and south poles on the north pole of the needle? On the south pole of the needle?

Figure 27.13 Magnetic field line pattern surrounding a bar magnet.



We can now use the alignment of a compass needle to map out **magnetic field lines** by placing a compass somewhere near a bar magnet, moving the compass a small distance in the direction of the needle, and repeating this procedure to trace out a line, as illustrated in **Figure 27.12**. At any location, the direction of the magnetic field $\vec{B}$ is tangent to the magnetic field line passing through the location. By convention:

Near a magnet, magnetic field lines point away from north poles and toward south poles.

A more complete diagram of the magnetic field lines outside a bar magnet is shown in **Figure 27.13**. Notice the similarity with an electric dipole field (Figure 24.2).

Magnetic field line patterns outside a magnet can be made visible by sprinkling iron filings on a piece of paper placed over the magnet. The little pieces of iron become magnetized and act like compass needles that align themselves along field lines (**Figure 27.14** on the next page).

The similarity between electric and magnetic field patterns suggests that we may be able to carry over many of the concepts we developed for electric fields into our study of magnetism. In particular, we can associate a magnetic field magnitude with the density of magnetic field lines:

At every location in a magnetic field line diagram, the magnitude of the magnetic field is proportional to the field line density at that location.

In the field pattern shown in **Figure 27.15*a*** on the next page, for example, the magnitude of the magnetic field is greatest

CONCEPTS

Figure 27.12 A magnetic field line can be traced by moving a compass small distances in the direction in which its needle points.

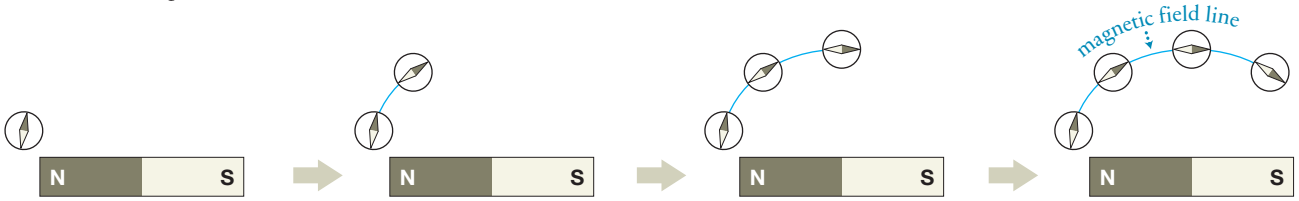
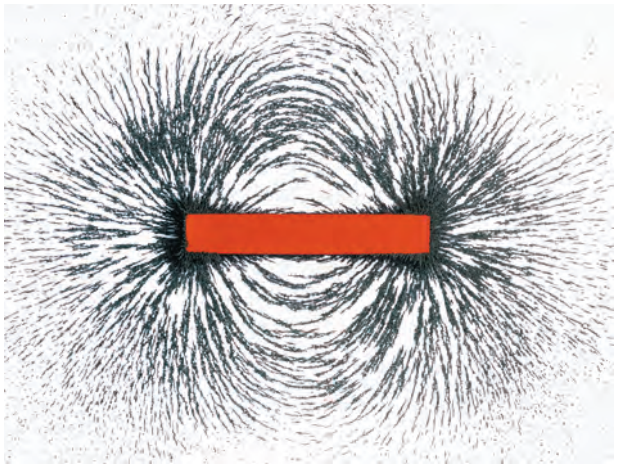


Figure 27.14 The magnetic field line pattern outside a bar magnet made visible by sprinkled iron filings around the magnet.

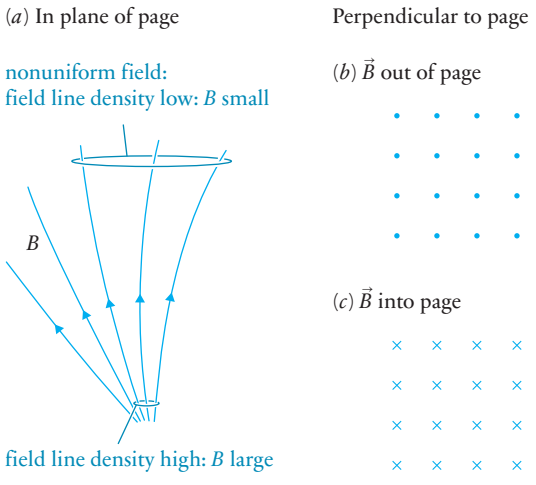


near the bottom and smallest near the top. Occasionally we shall work with magnetic fields that are perpendicular to the plane of the drawing, and so we need to be able to represent them. Figure 27.15*b* and *c* show conventions for representing field lines perpendicular to the page.

In electrostatics, the concept of field line flux led to Gauss's law, and so we may ask ourselves if there is an analogous law for magnetism. The answer is *yes*, but the result is somewhat different from that in electrostatics. To see why, complete this checkpoint.

27.6 (*a*) Consider a single elementary magnet inside a closed surface. Given that elementary magnets are particles without spatial extent, is the magnetic field line flux through the closed surface positive, negative, or zero? (*b*) Does adding a second elementary magnet inside the closed surface change your answer? (*c*) Consider a bar magnet inside a closed surface. Is the magnetic field line flux through the closed surface positive, negative, or zero? (*d*) Does your answer change if the closed surface cuts through the magnet?

Figure 27.15 Conventions for representing a magnetic field.



The fundamental point of Checkpoint 27.6 is that each field line leaving a closed surface that encloses an elementary magnet reenters the surface somewhere else, so the magnetic field line flux due to an elementary magnet inside a closed surface is zero. Generalizing this statement to a collection of many elementary magnets, we say:

The magnetic field line flux through a closed surface is always zero.

In electrostatics, electric field lines originate or terminate on electrical charge. Because of the absence of magnetic monopoles, however, we have no "magnetic charge" on which magnetic field lines could originate or terminate. Magnetic field lines must therefore always form loops that close on themselves. The statement that the magnetic field line flux through a closed surface is always zero is therefore a direct consequence of the absence of magnetic monopoles.

27.7 What is the direction of the magnetic field lines *inside* the bar magnet of Figure 27.13?

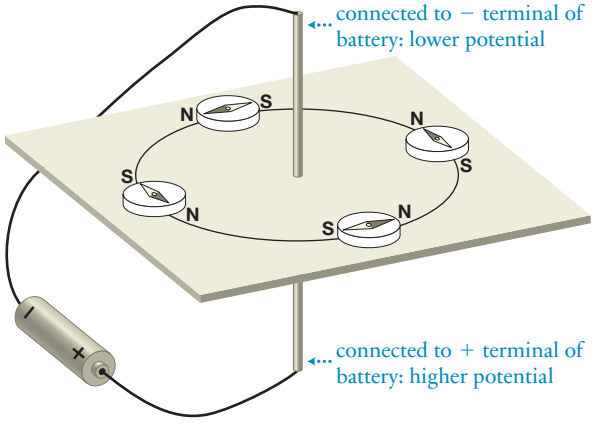
27.3 Charge flow and magnetism

The first indication of a connection between electricity and magnetism came in 1820, when the Danish physicist Hans Christian Ørsted discovered that a flow of charge carriers deflects the needle of a compass. This effect is illustrated in **Figure 27.16**, which shows battery terminals connected to a conducting rod.

The top of the rod in Figure 27.16 is connected to the negative battery terminal, and the bottom is connected to the positive terminal. As we saw in Chapter 26, this means that the bottom of the rod is at a higher potential than the top and an upward-pointing electric field arises inside the rod.*

*In Section 24.5 we concluded that the electric field inside a conductor in electrostatic equilibrium is zero. The rod in Figure 27.16, however, is not in electrostatic equilibrium because the battery maintains a potential difference across it, causing free charge carriers in the rod to move.

Figure 27.16 A flow of charge carriers through a conducting rod causes a circular alignment of compass needles.



CONCEPTS

The upward electric field in the rod exerts a downward electric force on the negatively charged electrons in the rod, causing a flow of charge carriers, called **current,** through the rod. The battery maintains a constant potential difference across the rod by adding electrons at the top of the rod and removing them at the bottom. Consequently, the battery maintains a constant current through the rod.

27.8 (*a*) Is the rod in Figure 27.16 electrically charged while connected to the battery? (*b*) Is there an electric field due to the rod at the positions of the compasses in Figure 27.16?

When the electrons flow through the rod in Figure 27.16, the compass needles align themselves in a circular pattern around the rod. This observation leads to a conclusion:

A flow of charged particles causes a magnetic field.

If the direction of the flow is reversed by reversing the battery, the compass needles turn around until they point in the opposite direction. If the flow of charge carriers is stopped, the needles point toward Earth's geographic North Pole, as usual.

27.9 Sketch the magnetic field line pattern in the horizontal plane around the rod in Figure 27.16.

It is important to note that a flow of positive charge carriers in one direction is equivalent to a flow of negative charge carriers in the opposite direction. In **Figure 27.17**, for example, the charge on the right increases and the charge on the left decreases regardless of which flow occurs. Experiments also show that the two types of flow are equivalent in terms of the magnetic field:

Figure 27.17 A flow of positive charge carriers in one direction is equivalent to a flow of negative charge carriers in the other direction.

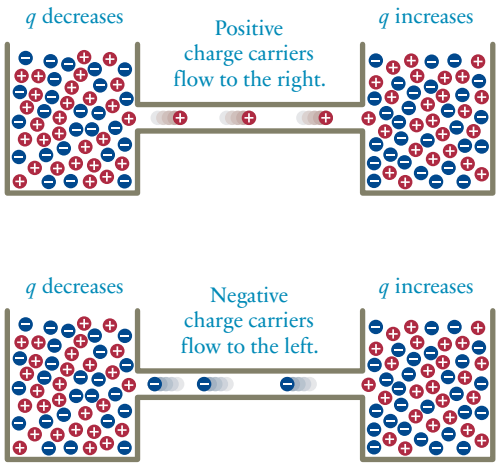
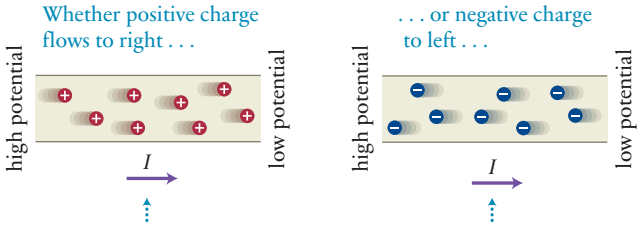


Figure 27.18 By definition, current has the direction in which positive charge carriers would flow, even if it is actually carried by negative charge carriers moving in the opposite direction.



. . . we say *direction of current* is to right (from high to low potential).

A flow of positive charge in one direction produces the same magnetic field as an equal flow of negative charge in the opposite direction.

Therefore, regardless of the actual movement of charge carriers through a current-carrying wire, we shall always denote the direction of decreasing potential (that is, the direction in which positive charge carriers would flow) with an arrow labeled with the symbol for current (I) next to the wire, as shown in **Figure 27.18**. We'll call this direction the *direction of current*.

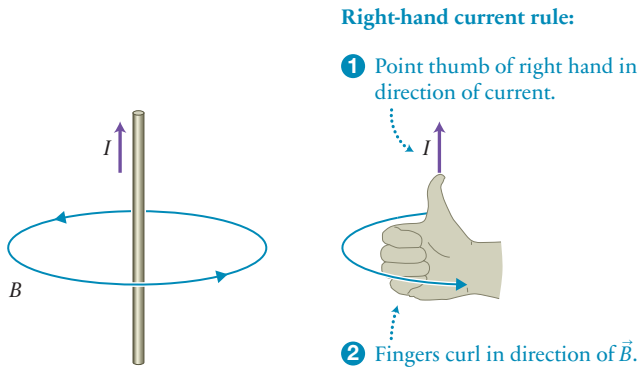
We can now connect the direction of a magnetic field to that of current. **Figure 27.19** shows a single circular magnetic field line for the current-carrying rod from Figure 27.16. As seen from the top, the magnetic field curls in a counterclockwise direction around the rod. The downward flow of negative electrons through the rod corresponds to an upward current, and so the direction of the magnetic field is connected to that of the current by the *right-hand current rule*:

If you point the thumb of your right hand in the direction of current, your fingers curl in the direction of the magnetic field produced by that current.

27.10 Can you replace the current-carrying rod of Figure 27.16 by a magnet and get the same magnetic field?

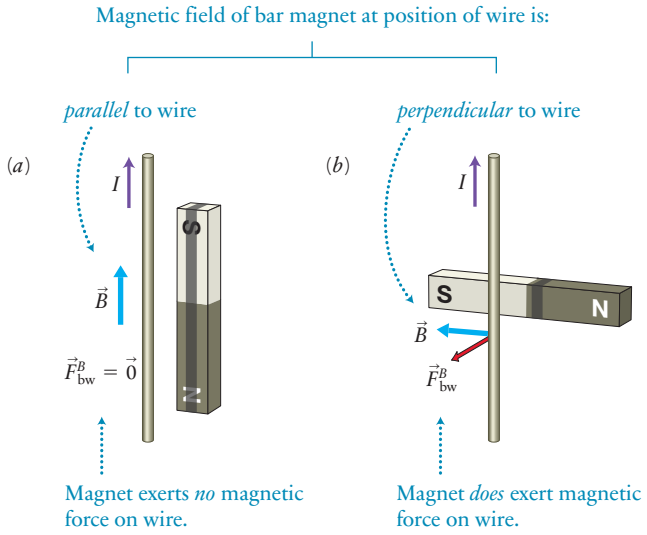
The relationship between the magnetic field produced by a magnet and that produced by a current-carrying wire

Figure 27.19 The right-hand current rule relates the direction of a current to the direction of the resulting magnetic field.



CONCEPTS

Figure 27.20 The magnetic force exerted by a bar magnet on a current-carrying wire depends on the magnet's orientation.



is not immediately obvious. We shall study this relationship in more detail in the next chapter. Before doing so, however, we still need to answer the question: How does a bar magnet interact with a straight current-carrying wire? The answer, which follows from the experiment illustrated in **Figure 27.20**, describes an interaction that is very different from all the interactions we have encountered so far. When a bar magnet is held parallel to a current-carrying wire, the magnetic field of the magnet at the position of the wire is parallel to the wire (Figure 27.20*a*). In this position, the magnet exerts *no force* on the wire! When we rotate the magnet 90°, so that its long axis is perpendicular to the long axis of the wire, the magnet exerts a repulsive force on the wire (Figure 27.20*b*). The most striking aspect of this interaction is that the direction of the magnetic force is *perpendicular* to the magnetic field.

CONCEPTS

27.11 Use Figure 27.20 to determine the direction of the magnetic force exerted by the magnet on the wire when the magnet is in the orientations shown (*a*) in **Figure 27.21*a*** and (*b*) in Figure 27.21*b*. (Hint: First determine the direction of the magnetic field due to the magnet at the wire; then use the observations from Figure 27.20 to determine the direction of the magnetic force.)

Figure 27.21 Checkpoint 27.11.

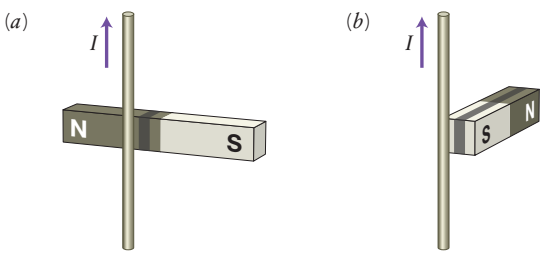
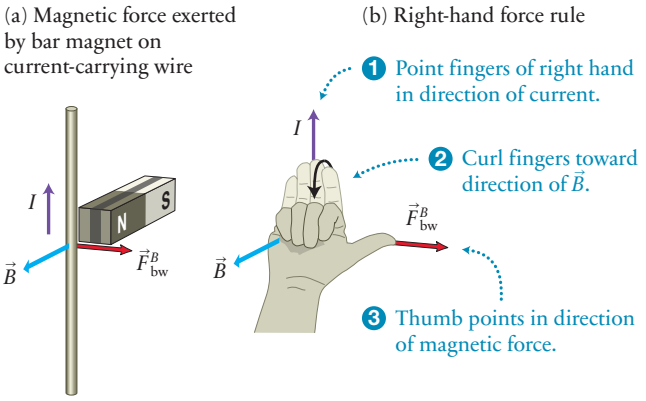


Figure 27.22 When a magnet exerts a force on a current-carrying wire, the right-hand force rule relates the direction of the force to that of the current and the magnet's magnetic field.



The surprising *sideways* force exerted by the bar magnet on the wire in Figure 27.21*b* (see Checkpoint 27.11) is a result of the fact that the magnetic force exerted by a bar magnet on a current-carrying wire is always at right angles to both the magnetic field and the direction of current. This observation suggests that these three directions can be connected by a right-hand rule. **Figure 27.22** shows that if you orient the fingers of your right hand in the direction of current in such a way that you can curl them so that the fingertips point in the direction of the magnetic field, then your right thumb points in the direction of the magnetic force exerted on the wire. Take a minute to verify the direction of the force in Figure 27.20*b* using this procedure, commonly called the *right-hand force rule*:*

> **The direction of the magnetic force exerted by a magnetic field on a current-carrying wire is given by the direction of the right-hand thumb when the fingers of that hand are placed along the direction of current in such a way that they can be curled toward the magnetic field.**

Table 27.1 summarizes the two right-hand rules we have encountered so far.

Table 27.1 Right-hand rules in magnetism

Right-hand rule	Thumb points along	Fingers curl
current rule	current	along *B*-field
force rule	magnetic force	from current to *B*-field

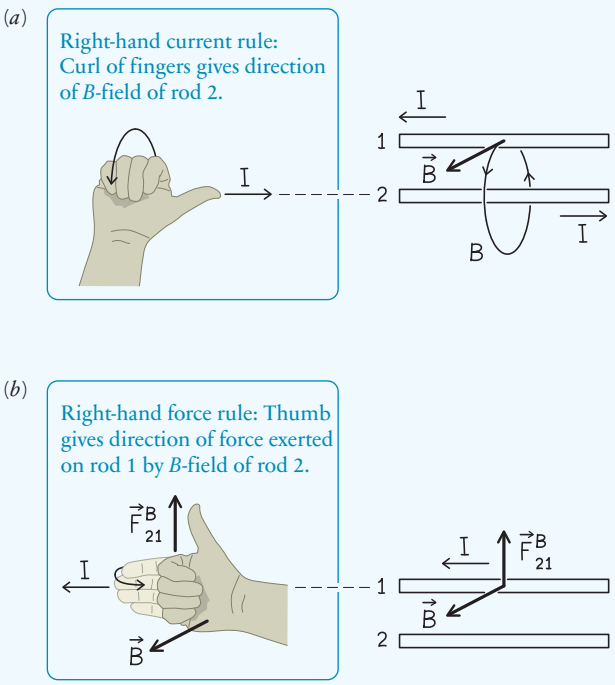
*There is nothing magical about right hands. You could use your left hand and change the rule to read, "The direction . . . given by the direction of the left-hand thumb when the fingers of that hand are placed along the direction of the magnetic field in such a way that they can be curled toward the direction of current." The same applies to determining the direction of a magnetic field surrounding a current-carrying rod. In Figure 27.19, for instance, orienting your left hand so that the thumb points in the direction *opposite* the direction of current automatically curls the fingers in the direction of the magnetic field. A consistent convention is what is important.

EXAMPLE 27.1 Current-carrying rods

Two parallel rods carry currents in opposite directions. Determine the direction of the magnetic force exerted by each rod on the other rod.

1 GETTING STARTED I begin by making a sketch of the rods, labeling them 1 and 2 and showing the currents in opposite directions (**Figure 27.23**). The current through rod 2 creates a magnetic field at the location of rod 1. This magnetic field exerts a force $\vec{F}_{21}^{B}$ on rod 1. Likewise, the magnetic field created at rod 2 by the current through rod 1 exerts a force $\vec{F}_{12}^{B}$ on rod 2.

Figure 27.23



2 DEVISE PLAN To determine the direction of each force, I need to apply the right-hand force rule. To apply this rule, I need to know the direction of the magnetic field created by each rod at the location of the other. I can determine these directions using the right-hand current rule.

3 EXECUTE PLAN I begin with the right-hand current rule to determine the direction of the magnetic field due to the current through rod 2 at the location of rod 1. When I align my right thumb with the direction of current through rod 2, my fingers curl out of the page (Figure 27.23*a*), telling me that this is the direction of the magnetic field due to rod 2. Next I apply the right-hand force rule to determine the direction of the magnetic force $\vec{F}_{21}^{B}$ exerted by rod 2 on rod 1. I align the fingers of my right hand with the current through rod 1 in such a way that they can curl toward $\vec{B}$ (Figure 27.23*b*). My right thumb now points up in the plane of the page, telling me that this is the direction of $\vec{F}_{21}^{B}$. ✔

To determine the direction of the magnetic force $\vec{F}_{12}^{B}$ exerted by rod 1 on rod 2, I again begin with the right-hand current rule, which tells me that, at the location of rod 2, the magnetic field due to rod 1 is directed out of the page. Applying the force rule, I place my fingers along the direction of I_2 and see that curling them in the direction of $\vec{B}$ makes my thumb point downward in the plane of the page, so this is the direction of $\vec{F}_{12}^{B}$. ✔

The force $\vec{F}_{21}^{B}$ points upward, the force $\vec{F}_{12}^{B}$ points downward, and therefore the rods repel each other.

4 EVALUATE RESULT It makes sense that the two rods exert forces on each other that point in opposite directions because the forces $\vec{F}_{12}^{B}$ and $\vec{F}_{21}^{B}$ form an interaction pair.

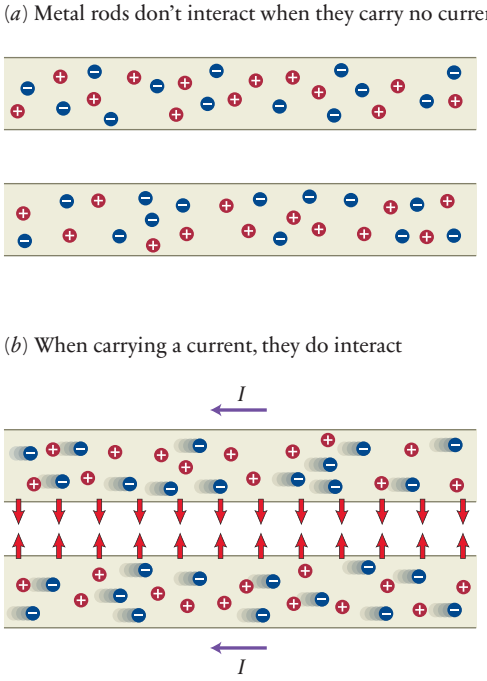


27.12 Determine the directions of the forces exerted by two parallel rods with currents in the same direction.

27.4 Magnetism and relativity

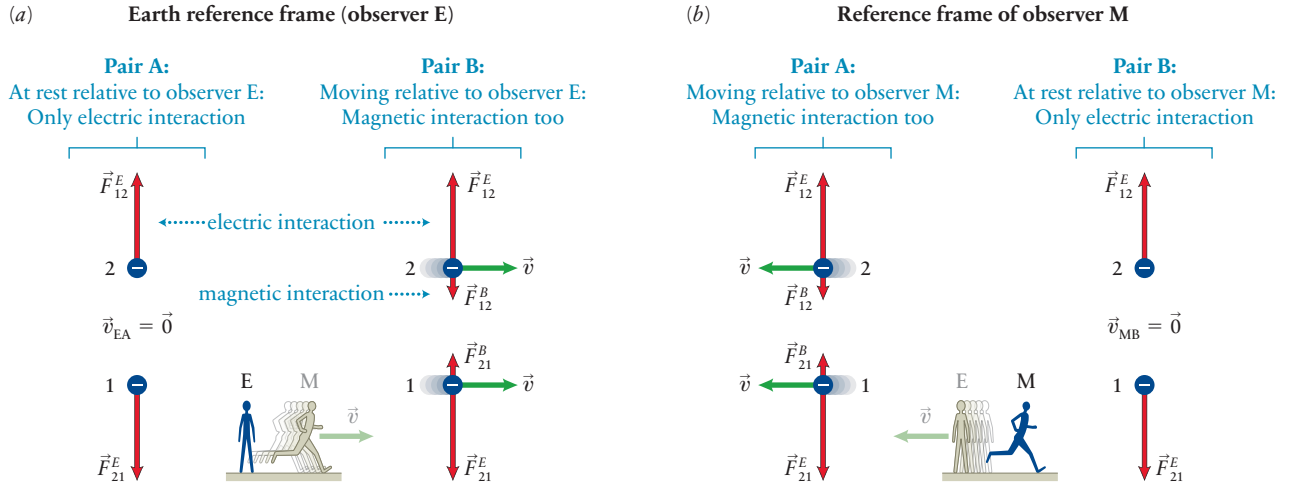
The magnetic interaction between two current-carrying rods is baffling. How can any interaction between the rods depend on the motion of the charge carriers in them? **Figure 27.24*a*** shows a schematic view of the positive and negative charge carriers that make up two parallel metal rods. The electrons in the two rods repel one another, but this repulsion is perfectly balanced by the attraction between the electrons in one rod and the positively charged ions in the other rod. Once the electrons are set in motion, however, this balance appears to be disturbed, and, depending on the signs of the

Figure 27.24 Schematic view of the interaction between two current-carrying rods.



CONCEPTS

Figure 27.25 A relativistic view of the interaction between two charged particles.



currents, the rods either attract (Figure 27.24*b*) or repel each other.

To analyze this interaction, consider the simpler situation of two pairs of electrons. Relative to Earth, pair A is at rest and pair B is moving at constant velocity $\vec{v}$ (**Figure 27.25*a***). To observer E, who is also at rest relative to Earth, the interaction between the two electrons of pair A is completely described by Coulomb's law: As we saw in Chapter 22, the electrons repel each other due to an electric interaction. According to observer E, the interaction between the electrons of pair B is *different* from that between the pair A electrons. In addition to the Coulomb repulsion, observer E perceives a magnetic interaction between the electrons of pair B because the moving electrons constitute a current.

The conclusions drawn by an observer M moving along with pair B are startlingly different (Figure 27.25*b*). To this observer, pair B is at rest and pair A is moving to the left at constant velocity. Consequently, to M the interaction between the electrons of pair B is purely electric, while that between the electrons of pair A must have a magnetic component as well. This odd conclusion is a direct consequence of the fact that magnetism is motion-dependent and therefore, like velocity, must be relative:

The observed interaction between charge carriers depends on their motion relative to the observer: The interaction can be purely electric, purely magnetic, or a combination of the two.

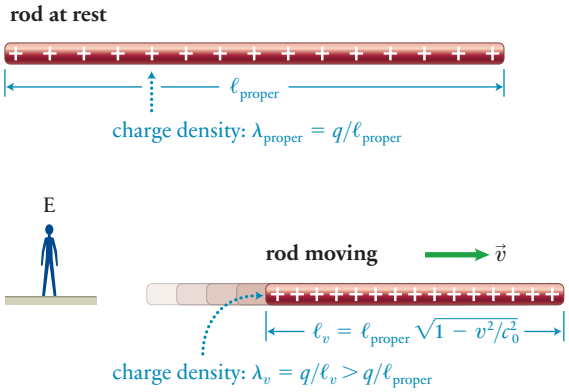
At first sight, this statement appears to be a violation of the principle of relativity. If reference frame M moves at constant velocity relative to reference frame E, then both observers should agree on their observations (see Section 6.3). The resolution of this problem lies in the fact that the magnetic interaction between two moving charged carriers is a direct consequence of special relativity: *Magnetism is a relativistic effect.*

As explained in Chapter 14, relativistic effects tend to be so extraordinarily small that they cannot be observed under most ordinary conditions. So why is it that we can feel this relativistic effect every time we stick a note to a refrigerator with a magnet? To see why, let us return to the two current-carrying rods in Figure 27.24. Two 1-m-long rods, each 0.5 mm in diameter, contain an immense number of free electrons—more than 10^{22} of them. To appreciate how enormous this number is, imagine that these electrons were not accompanied by an equal number of positively charged ions. The force with which two such collections of unbalanced electrons separated by a distance of 5 mm would repel each other is a phenomenal 10^{19} N—enough to lift up thousands of average-sized mountains. The fact that two electrically neutral metal rods do not exert any force on each other shows how incredibly accurate is the balance between positive and negative charge carriers in matter. In contrast, if we run large currents through the rods—by applying and maintaining a potential difference of, say, several volts across the length of each rod—then the magnetic force exerted by each rod on the other is a mere 10^{-2} N. This small but measurable magnetic force is 10^{21} times smaller than the electric repulsion between the electrons in the rods! Thus, if magnetism is indeed a relativistic effect (see below and Section 27.8 for details), we can now understand why we can so readily observe it. First, the electric force is incredibly large, and so even a small relativistic correction becomes measurable. Second, because electrical charge is so well balanced in all matter, electric forces are balanced, leaving only the small "magnetic" correction. In a sense, therefore, magnetism provides the most direct observation of a relativistic effect!

Let us now try to understand qualitatively how special relativity requires a magnetic interaction between charge carriers in motion. The treatment that follows

CONCEPTS

Figure 27.26 Because a rod appears shorter when it is moving relative to the observer, its charge density appears larger than when the rod is at rest.



relies on very little knowledge of special relativity—only the concept of length contraction is needed (see Section 14.6): When an object moves relative to an observer, the observer measures the length of the object in the direction of its motion to be shorter than the proper length of the object.

A direct result of length contraction is that the charge density of an object depends on its motion relative to the observer. Consider, for example, a rod carrying a positive charge q. At rest in the Earth reference frame, the rod has length ℓ_{proper} and its "proper" charge density is $\lambda_{\text{proper}} = q/\ell_{\text{proper}}$. If the rod is set in motion along its lengthwise axis, its length measured by an observer at rest is smaller than ℓ_{proper} (**Figure 27.26**). Because the charge on the rod is still q, the observer at rest sees a greater charge density on the moving rod.*

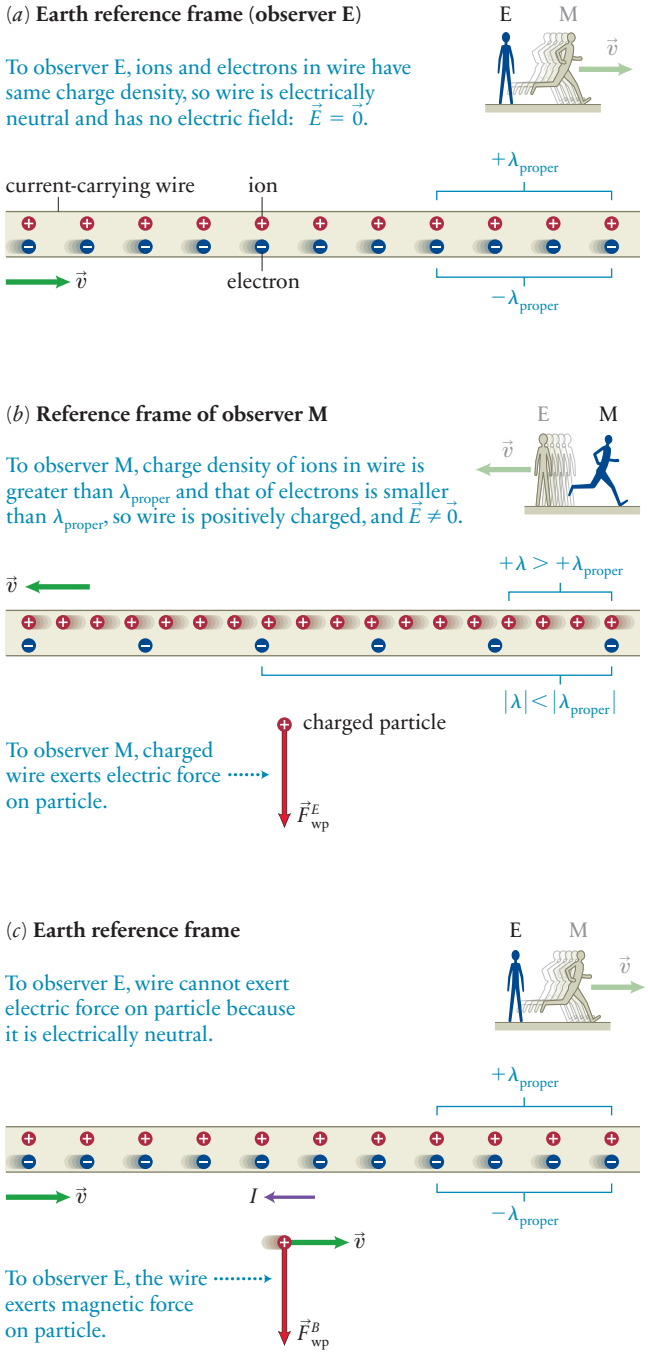


27.13 Consider the two identical rods in Figure 27.26, one moving and the other at rest relative to observer E, at rest in the Earth reference frame. Suppose a second observer M moves along with the moving rod. (*a*) Which rod has the greater charge density according to observer M? (*b*) Suppose the charge on the rod moving in the Earth reference frame is adjusted so that its charge *density* as seen by observer E is the same as the charge density of the rod at rest in the Earth reference frame. Is the charge density on each rod as seen by observer M greater than, equal to, or smaller than $\lambda_{\text{proper}} = q/\ell_{\text{proper}}$?

Understanding the changes in charge density that occur when charge carriers move relative to the observer is the key point for the remainder of this section. Thus

*It is not immediately obvious that the charge of a system is not affected by the motion of the charge carriers. Experiments, however, show that charge is, indeed, an invariant (that is, its value does not depend on the choice of reference frame).

Figure 27.27 Observers E (at rest in the Earth reference frame) and M (moving in the Earth reference frame) observe a current-carrying wire.



if you did not complete the preceding checkpoint, please go back and do it now before reading on.

Let us now apply these ideas to the interaction between a current-carrying wire and a charged particle. **Figure 27.27***a* schematically shows a current-carrying wire consisting of fixed positively charged ions (red) and negatively charged electrons (blue) traveling to the right at velocity $\vec{v}$. (For clarity, the ions and electrons are drawn side by side rather than intermingled as they are in a real conductor.)

CONCEPTS

To an observer E at rest relative to the ions, the wire does not carry a surplus charge, and so any length of wire contains the same number of ions and electrons. That is, to E the linear charge density of the ions, λ_{proper}, is equal in magnitude to that of the electrons, $-\lambda_{\text{proper}}$. Consequently, there is no electric field outside the wire, and the wire cannot exert an electric force on a charged particle placed near the wire.

Suppose, however, that a positively charged particle moves alongside the wire—for simplicity, we let it move at the same velocity $\vec{v}$ relative to the wire as the electrons in the wire. First we consider the particle and the wire from the point of view of an observer M moving along with the particle. Because the particle is at rest relative to M, it experiences no magnetic force according to this observer. From M's point of view, the electrons in the wire are at rest and the ions move at velocity $\vec{v}$ to the left (Figure 27.27*b*). In addition, as you saw in Checkpoint 27.13, the charge densities seen by M are different from those seen by E. To M, the positive ions are closer together because they are moving, which means their charge density is greater than λ_{proper}. The oppositely charged electrons, on the other hand, are at rest relative to M and so farther apart. According to observer M the magnitude of their linear charge density therefore must be smaller than that observed by E relative to whom they are moving.

The different charge densities mean that although the wire appears electrically neutral to E, it cannot also appear neutral to M. According to M, the magnitude of the electron density is smaller than λ_{proper}, and the magnitude of the ion density is greater than λ_{proper}. Thus, the wire appears positively charged to observer M. Observer M therefore sees a downward electric field due to the wire at the location of the charged particle, with the wire and the particle repelling each other.

In the reference frame of the fixed ions, a very different picture emerges. Observer E, seeing no electric field (Figure 27.27*a*), cannot attribute the repulsion between the wire and the particle to an electric interaction. Instead, E attributes this repulsion to a magnetic interaction between two currents, one in the wire and the other caused by the moving charged particle (Figure 27.27*c*). As we shall see in Section 27.8, this magnetic interaction can be completely accounted for by the Coulomb interaction and special relativity, and in principle any other magnetic interaction can also be explained this way. However, because transforming back and forth from one moving reference frame to another is cumbersome, it is easier to develop a separate treatment for magnetism that does not require reference-frame transformations and that ignores any relativistic effects. It is important to keep in mind, however, that magnetic and electric interactions are two aspects of one *electromagnetic* interaction, with magnetism being a relativistic correction to the electric interaction.

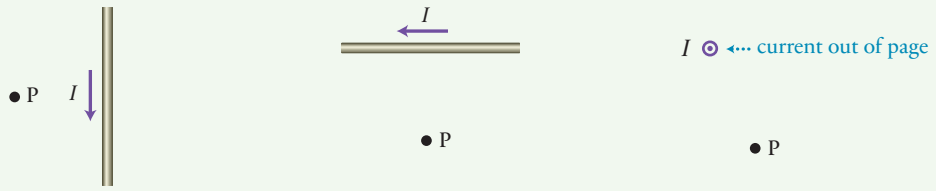
CONCEPTS

27.14 What is the direction of current through the rod in (*a*) Figure 27.27*b* and (*b*) Figure 27.27*c*? (*c*) Is the direction of the force in Figure 27.27*c* in agreement with what we learned about the interaction of two parallel current-carrying wires in Section 27.3? (*d*) If the particle moving alongside the wire in Figure 27.27 carries a negative charge, is the force exerted by the wire on the particle attractive or repulsive? (*e*) Is this direction consistent with the direction of the forces exerted on each other by two parallel current-carrying wires?

Self-quiz

1. A compass sits on a table with its needle pointing to Earth's North Pole. A bar magnet with its long axis oriented along an east-west line is brought toward the compass from the east. If the needle turns clockwise, which pole of the bar magnet is nearer the compass?
2. Draw the magnetic field lines associated with the magnets in Figures 27.1*b–d* (both outside and inside the magnets). Assume that the pole on the left of each magnet is the north pole.
3. For each situation shown in **Figure 27.28**, apply the appropriate right-hand rule to determine the direction at position P of the magnetic field generated by the current-carrying wire.

Figure 27.28

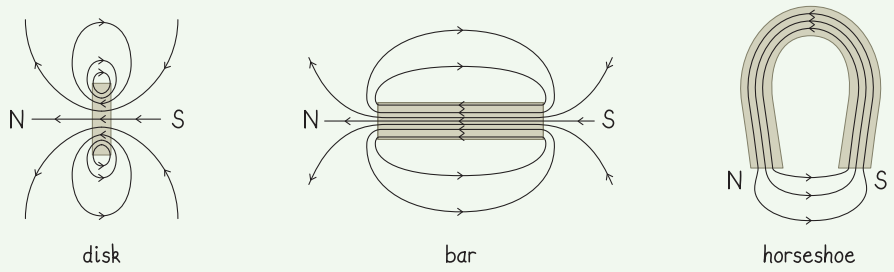


4. For each situation in Figures 27.20*a*, 27.20*b*, and 27.21*a*, reverse the polarity of the magnet and then apply the appropriate right-hand rule to determine the direction of the force exerted by the bar magnet on the current-carrying wire.

Answers

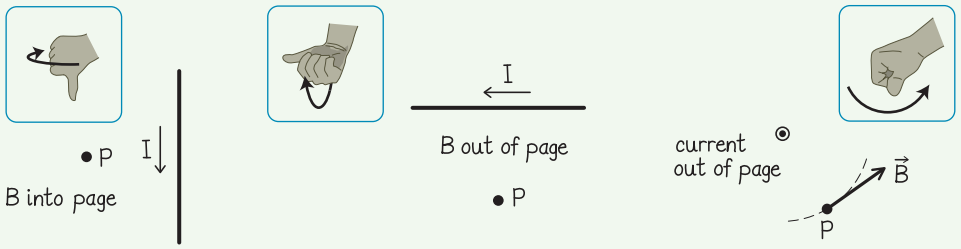
1. South pole. The clockwise rotation means the needle tip (which is a north pole by definition) moves toward the bar magnet. Because opposite poles attract, the bar magnet's south pole must be the closer one. (Remember that Earth's geographic North Pole is a magnetic south pole and so attracts the north pole of any compass needle.)
2. See **Figure 27.29**.

Figure 27.29



3. See **Figure 27.30**. (See Checkpoint 27.7 if you do not understand this answer.)

Figure 27.30



4. Reversing the magnet polarity reverses the direction of the force exerted on the wire, unless the force is zero, in which case it remains zero.

27.5 Current and magnetism

As we saw in the first part of this chapter, currents exert magnetic forces on one another. That is, a current creates a magnetic field, and magnetic fields exert forces on currents. Let us therefore begin this part of the chapter by introducing a quantitative definition of current. As mentioned in Section 27.3, **current** I is the rate at which charged particles cross a section of a conductor in a given direction. For a constant current, we have

$$I \equiv \frac{q}{\Delta t} \quad \text{(constant current)}, \tag{27.1}$$

where q is the quantity of charge passing a given position in a time interval Δt. If the current is not constant, we evaluate the flow of charged particles over infinitesimally small time intervals, yielding

$$I \equiv \frac{dq}{dt}, \tag{27.2}$$

where dq is the infinitesimal quantity of charge crossing a given section of a conductor in an infinitesimally small time interval dt. Note that I can be positive or negative; by definition it is positive in the direction from high to low potential because that is the direction in which positive charge carriers flow (which we defined earlier as the direction of current).

The SI unit of current is the **ampere** (A). This base unit is now defined in terms of the precise value of the charge of the electron—that is, the elementary electric charge. The elementary charge is equal to exactly $1.602176634 \cdot 10^{-19}$ ampere-seconds (A · s). As we saw in Chapter 22, the coulomb is derived from the ampere: 1 C corresponds to the quantity of charge transported by a current of 1 A through a chosen section in a time interval of 1 s, or

$$1\ \text{C} \equiv 1\ \text{A} \cdot \text{s}. \tag{27.3}$$

Even though a charge of 1 C is extremely large by ordinary standards, currents of several amperes are quite common. Simple devices, which we shall discuss in Chapter 31, allow us to measure the current through a conductor quite readily. In the remainder of this chapter, we shall discuss constant currents—that is, currents whose magnitude does not change with time.

Experiments show that when a straight wire carrying a current I is placed in a uniform external magnetic field,* the magnetic force $\vec{F}_{\text{w}}^{B}$ exerted by the magnetic field on the wire is proportional to the length ℓ of wire in the magnetic field and to the magnitude of the current I. The force also depends on the angle θ between the direction of the current through the wire and the magnetic field (**Figure 27.31**). When the wire is parallel to the magnetic field ($\theta = 0$), the magnetic force exerted on the wire is zero. When the wire and the magnetic field are perpendicular to each other ($\theta = 90°$), the force is maximum:

$$F_{\text{w, max}}^{B} = |I|\ell B \quad \begin{matrix}\text{(straight wire perpendicular} \\ \text{to uniform magnetic field).}\end{matrix} \tag{27.4}$$

Because I is a signed quantity, we need to put absolute-value symbols around it to indicate its magnitude.

Equation 27.4 defines the magnitude B of the magnetic field. If we measure the magnetic force exerted by a magnetic field on a wire of known length carrying a known current, we can determine B from Eq. 27.4:

$$B \equiv \frac{F_{\text{w, max}}^{B}}{|I|\ell} \quad \begin{matrix}\text{(straight wire perpendicular} \\ \text{to uniform magnetic field).}\end{matrix} \tag{27.5}$$

*This magnetic field does not include the magnetic field of the current-carrying wire.

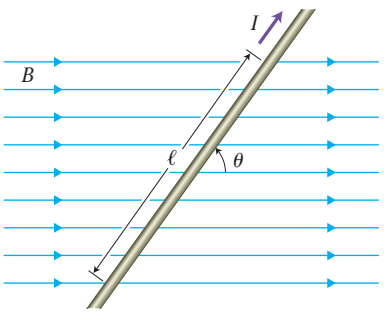


Figure 27.31 A current-carrying wire in an external magnetic field (that is, a magnetic field created by an object other than the wire).

QUANTITATIVE TOOLS

It follows from this expression that the magnetic field has SI units of $N/(A \cdot m)$, a derived unit called the **tesla** (in honor of engineer and inventor Nikola Tesla):

$$1\ T \equiv 1\ N/(A \cdot m) = 1\ kg/(s^2 \cdot A). \quad (27.6)$$

A magnetic field of 1 T is relatively large. Earth's magnetic field at Earth's surface varies between 3×10^{-5} T and 6×10^{-5} T. Table 27.2 provides some examples of the magnitudes of various magnetic fields.

At intermediate angles θ, the magnitude of the magnetic force exerted on the wire is proportional to $\sin\theta$, and so for an arbitrary angle θ between 0 and 180°, the magnitude of the magnetic force exerted on the wire is

$$F_w^B = |I|\ell B \sin\theta \quad (0 < \theta < 180°). \quad (27.7)$$

As we saw in Figure 27.22, the direction of the magnetic force exerted by a magnet on a current-carrying wire is always perpendicular to both the direction of the current I and the magnetic field $\vec{B}$, and is given by the right-hand force rule. In Figure 27.31, for example, the force is directed into the plane of the page. (If you line up the fingers of your right hand with the direction of current and then curl them, over the smallest angle, toward the direction of the magnetic field, your thumb points into the page.) If we define a vector $I\vec{\ell}$, whose magnitude is given by the product of the magnitude of the current and the length ℓ of the wire, and whose direction is given by the direction of current through the wire, we can write the **magnetic force** in Eq. 27.7 as the vector product of two vectors (see Section 12.8):

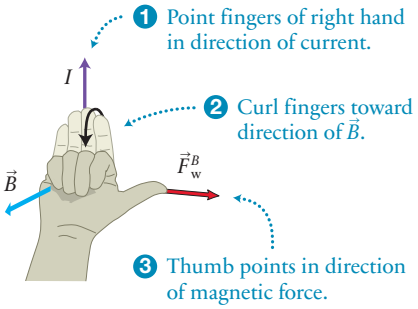
$$\vec{F}_w^B = I\vec{\ell} \times \vec{B} \quad (\text{straight wire in uniform magnetic field}). \quad (27.8)$$

This vector product represents a force that has the magnitude given in Eq. 27.7. The direction of $\vec{F}_w^B$ is always perpendicular to both $I\vec{\ell}$ (the direction of current) and $\vec{B}$ and is obtained by curling the fingers of the right hand from the direction of current to $\vec{B}$ (Figure 27.32).

Table 27.2 Magnetic fields

Object	B (T)
Earth's surface	5×10^{-5}
small bar magnet	0.01
neodymium magnet	0.2
laboratory magnet	10
neutron star surface	10^8

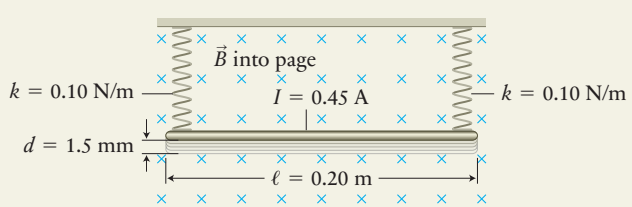
Figure 27.32 The right-hand rule for the direction of the vector product $I\vec{\ell} \times \vec{B}$.



EXAMPLE 27.2 Magnetic field meter

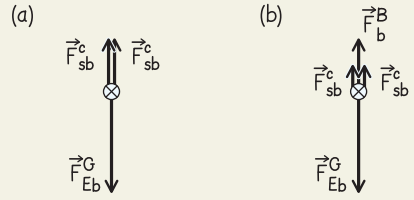
A metal bar 0.20 m long is suspended from two springs, each with spring constant $k = 0.10$ N/m, and the bar is in an external magnetic field directed perpendicular to the bar length (Figure 27.33). With a current of 0.45 A in the bar, the bar rises a distance $d = 1.5$ mm. (*a*) In what direction is the current? (*b*) What is the magnitude B of the external magnetic field?

Figure 27.33 Example 27.2.



❶ **GETTING STARTED** I begin by making a free-body diagram for the bar with and without the current through the bar. Without the current, the bar is subject to two upward forces exerted by the springs and a downward force of gravity (Figure 27.34*a*). With the current turned on, an upward magnetic force lifts the bar, and so, according to Hooke's law (see Section 8.9), the springs exert a smaller force on the bar (Figure 27.34*b*). The amount by which the bar rises determines by how much the force exerted by the springs is reduced, which in turn gives me the magnitude of the upward magnetic force exerted on the bar.

Figure 27.34



(Continued)

2 DEVISE PLAN I know the directions of the external magnetic field (into the page) and of the force (upward) exerted by that field, so I can determine the direction of the current from the right-hand force rule. Because the bar is straight and perpendicular to the external magnetic field, I can use Eq. 27.5 to relate the magnitude of the magnetic field to the magnitude of the magnetic force. I know both the current and the length of the bar but not the magnitude of the force exerted on the bar. I know, however, that this force magnitude is equal to the change in the magnitude of the force exerted by the springs, which I can calculate using Hooke's law (Eq. 8.20).

3 EXECUTE PLAN (*a*) The external magnetic field is into the plane of the page (Figure 27.33) and the magnetic force is directed upward. I therefore lay my right hand on top of Figure 27.33 with my thumb pointing up to represent the upward force. I can lay my hand on the page palm up or palm down. In the palm-up position, curling my fingers draws the tips away from the direction of $\vec{B}$. In the palm-down position, curling them draws the tips toward the direction of $\vec{B}$. Therefore that's the position I want, telling me the current is to the right, in the direction my fingers point before I curl them. ✔

(*b*) From Eq. 8.20, I know that raising the bar a distance d reduces the spring force by an amount kd, and so the magnitude of the magnetic force is $F_b^B = 2kd$. Substituting this result into Eq. 27.5 gives

$$B = \frac{2kd}{|I|\ell} = \frac{2(0.10\ \text{N/m})(1.5\times10^{-3}\ \text{m})}{(0.45\ \text{A})(0.20\ \text{m})} = 3.3\times10^{-3}\ \text{T}. \checkmark$$

4 EVALUATE RESULT The magnitude of the magnetic field I obtain is about 100 times greater than Earth's magnetic field and of the same order of magnitude as that of a bar magnet and therefore not unreasonable.



27.15 Suppose the charge carriers flowing through the horizontal bar in Example 27.2 are negatively charged. In which direction must they flow so that the magnetic force exerted on the wire is still directed upward?

27.6 Magnetic flux

From the field line picture for electrostatics, we obtained Gauss's law, which states that the electric flux Φ_E through any closed surface is proportional to the enclosed charge. Gauss's law allows us to determine the electric field of certain symmetrical charge distributions with great ease.

Because magnetic fields can also be described by field lines, we can follow a similar treatment for magnetism. Consider a surface of area A in a uniform magnetic field $\vec{B}$ (**Figure 27.35**). If a line normal to the surface makes an angle θ with the field, we can define a **magnetic flux** in analogy to the electric flux defined in Eq. 24.2:

$$\Phi_B \equiv BA\cos\theta = \vec{B}\cdot\vec{A} \quad \text{(uniform magnetic field)}, \tag{27.9}$$

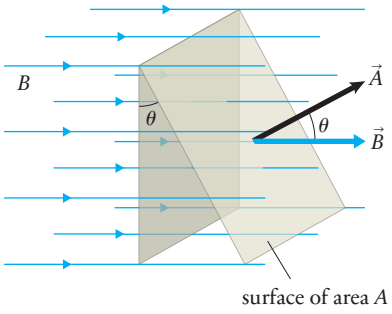
where $\vec{A}$ is an area vector (see Section 24.6) whose magnitude is equal to the area A of the surface and whose direction is normal to the surface. In the case of a closed surface, the direction of $\vec{A}$ is always chosen to be outward. Magnetic flux is a scalar, and as you can see from Eq. 27.9, magnetic flux has SI units of $\text{T}\cdot\text{m}^2$. This derived unit is given the name **weber** (Wb), in honor of the German physicist Wilhelm Weber:

$$1\ \text{Wb} \equiv 1\ \text{T}\cdot\text{m}^2 = 1\ \text{m}^2\cdot\text{kg}/(\text{s}^2\cdot\text{A}).$$

If the field is nonuniform or the surface is not flat, we follow the same procedure as for the electric flux and divide the surface into small surface elements, apply Eq. 27.9 to each surface element, and then sum the magnetic flux through all the elements. In the limit that the area of each element approaches zero, the sum is replaced by a surface integral (Eq. 24.4):

$$\Phi_B \equiv \int \vec{B}\cdot d\vec{A}, \tag{27.10}$$

Figure 27.35 The magnetic flux through a surface of area A is given by the scalar product of the area vector $\vec{A}$ and the magnetic field $\vec{B}$.



QUANTITATIVE TOOLS

where $d\vec{A}$ is the area vector of an infinitesimally small segment of the surface. The meaning of magnetic flux Φ_B is similar to that of electric flux. It is a quantitative measure of the number of magnetic field lines crossing the surface specified in the integration.

As we saw in Section 27.2, magnetic field lines always form loops. Thus, the magnetic flux through any closed surface must always be zero because every field line that exits the surface must enter it somewhere else:

$$\Phi_B = \oint \vec{B} \cdot d\vec{A} = 0. \tag{27.11}$$

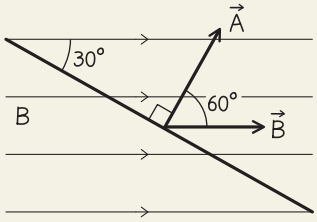
The zero reflects the fact that magnetic field lines always form loops or, equivalently, that there is no magnetic equivalent of an isolated charged particle. Equation 27.11 is called **Gauss's law for magnetism.** However, because of the zero result and because magnetic fields generally do not exhibit the same type of symmetry as electric fields, this expression does not allow us to determine magnetic fields in the same way that Gauss's law allows us to determine electric fields. As we shall see in Chapter 29, however, the magnetic flux is an important quantity when we consider changing magnetic fields.

EXAMPLE 27.3 Magnetic flux through a loop

A square loop 0.20 m on each side is placed in a uniform magnetic field of magnitude 0.50 T. The plane of the loop makes a 30° angle with the magnetic field. What is the magnetic flux through the loop?

❶ **GETTING STARTED** I begin by drawing a side view to visualize the situation (Figure 27.36). I have to calculate the magnetic flux through the flat surface defined by the loop.

Figure 27.36



❷ **DEVISE PLAN** Because the magnetic field is uniform, I can use Eq. 27.9 to calculate the magnetic flux through the loop.

❸ **EXECUTE PLAN** The area of the loop is $A = (0.20\ \text{m})^2 = 0.040\ \text{m}^2$. Because the loop makes a 30° angle with the magnetic field, the angle between the magnetic field and a normal to the plane of the loop is 60°. Therefore

$$\Phi_B = AB\cos\theta = (0.040\ \text{m}^2)(0.50\ \text{T})\cos(60°)$$
$$= 1.0 \times 10^{-2}\ \text{Wb}. \checkmark$$

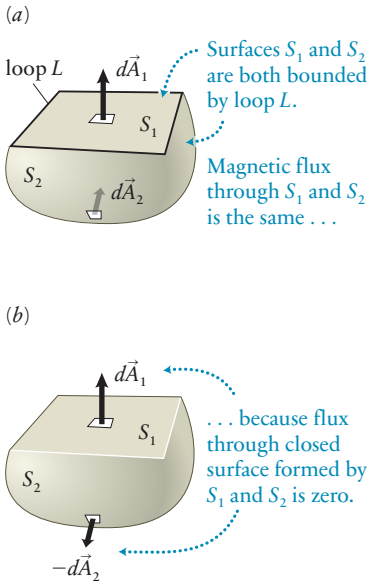
❹ **EVALUATE RESULT** I arbitrarily chose the area vector $\vec{A}$ to point upward and to the right. Had I chosen to point it downward and to the left, the angle between $\vec{A}$ and $\vec{B}$ would have been 120° and, given that $\cos 120° = -\frac{1}{2}$, I would have obtained $\Phi_B = -0.010$ Wb. The flat surface defined by the loop doesn't constitute a closed surface, however, and so there is no unique direction for $\vec{A}$. The sign of the magnetic flux is therefore not determined by the information given in the question. I have no means of evaluating the magnitude of the flux, so I carefully check my calculations one more time.

Because we can draw an infinite number of surfaces that are bounded by a loop, the expression "flux through the loop" in Example 27.3 appears open to interpretation. Figure 27.37*a* on the next page, for example, shows two surfaces S_1 and S_2 bounded by the same loop L. Which surface should we use when computing the magnetic flux? The answer is: *any* surface bounded by L yields the same magnetic flux. To see why, let us examine the magnetic fluxes through S_1 and S_2. In both cases we use a surface normal that points up through loop L.

The magnetic flux through S_2 is

$$\Phi_B = \int_{S_2} \vec{B} \cdot d\vec{A}_2. \tag{27.12}$$

Figure 27.37 Any surface bounded by loop L yields the same magnetic flux.



To see how this flux relates to that through S_1, consider the magnetic flux through the closed surface made up of S_1 and S_2 (Figure 27.37*b*). Because this surface is closed, we know the magnetic flux through it is zero (Eq. 27.11):

$$\oint_{S_1+S_2} \vec{B}\cdot d\vec{A} = \int_{S_1} \vec{B}\cdot d\vec{A} + \int_{S_2} \vec{B}\cdot d\vec{A} = 0, \tag{27.13}$$

where the area vector $d\vec{A}$ points *outward* from the closed surface. In terms of the surface normals $d\vec{A}_1$ and $d\vec{A}_2$ shown in Figure 27.37*a*, Eq. 27.13 becomes

$$\oint_{S_1+S_2} \vec{B}\cdot d\vec{A} = \int_{S_1} \vec{B}\cdot d\vec{A}_1 + \int_{S_2} \vec{B}\cdot(-d\vec{A}_2)$$

$$= \int_{S_1} \vec{B}\cdot d\vec{A}_1 - \int_{S_2} \vec{B}\cdot d\vec{A}_2. \tag{27.14}$$

The term $\int_{S_1} \vec{B}\cdot d\vec{A}_1$ represents the magnetic flux through the flat surface S_1; the term $\int_{S_2} \vec{B}\cdot d\vec{A}_2$ represents the magnetic flux through the curved surface S_2. Because the magnetic flux through the closed surface is zero (Eq. 27.11), the right side of Eq. 27.14 must also be zero, and so

$$\int_{S_1} \vec{B}\cdot d\vec{A}_1 = \int_{S_2} \vec{B}\cdot d\vec{A}_2. \tag{27.15}$$

27.16 A cube 1.0 m on each side is placed in a 1.0-T magnetic field with the field perpendicular to one surface of the cube. What are (*a*) the magnetic flux through the side through which the field enters the cube and (*b*) the magnetic flux through the entire surface of the cube?

27.7 Moving particles in electric and magnetic fields

The magnetic force exerted by a magnetic field on a straight current-carrying wire is really the sum of the magnetic forces exerted on many individual charge carriers moving through the wire. By examining how a current I through a conductor is related to the properties of the charge carriers causing that current, we can therefore deduce the magnetic force acting on a single charge carrier.

Consider a constant current caused by a flow of charge carriers, each carrying a quantity of charge q through a wire of cross-sectional area A (**Figure 27.38**). If the charge carriers flow at speed v, then in a time interval Δt each advances a distance $\ell = v\Delta t$. Thus all the charge carriers in the shaded volume $V = A\ell$ pass through a cross section of the wire in a time interval Δt. If the wire contains n charge carriers per unit volume, the shaded volume contains nV charge carriers and the charge flowing through the cross section is $Q = nVq = nA\ell q = nA(v\Delta t)q$. The current through the wire is thus

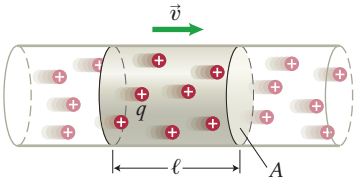
$$I \equiv \frac{Q}{\Delta t} = \frac{nA(v\Delta t)q}{\Delta t} = nAqv. \tag{27.16}$$

Substituting this result into Eq. 27.7, we can write for the magnitude of the magnetic force exerted on the current-carrying wire

$$F_w^B = |nA\ell qv|B\sin\theta = nA\ell|q|vB\sin\theta, \tag{27.17}$$

where θ is the angle between the velocity of the charge carriers and the magnetic field. Because n is the number of charge carriers per unit volume and $A\ell$ is the volume of a length ℓ of the wire, the quantity $N = nA\ell$ represents the number of charge carriers in a length ℓ of the wire. The magnetic force exerted on that length of the wire can thus be written

Figure 27.38 If the charge carriers in a straight current-carrying wire move at speed v, they advance a distance $\ell = v\Delta t$ in a time interval Δt. All the charge carriers in the shaded volume pass through the cross-sectional area A in that time interval.



QUANTITATIVE TOOLS

$F_w^B = N|q|vB \sin\theta$. Because there are N charge carriers in the length ℓ, the magnitude of the **magnetic force** exerted on a single particle carrying a charge q moving at velocity $\vec{v}$ is

$$F_p^B = |q|vB \sin\theta \tag{27.18}$$

or, in vector form,

$$\vec{F}_p^B = q\vec{v} \times \vec{B}. \tag{27.19}$$

Note that the vector $q\vec{v}$ always points in the direction of current—that is, from high potential to low potential—regardless of the sign of q. As we saw in Section 25.3, a negatively charged particle moves from low to high potential, so $\vec{v}$ points opposite the direction of current, but because $q < 0$, the vector $q\vec{v}$ points in the direction of current.

This is the fundamental expression for the magnetic force acting on a moving charge carrier. Even though we derived Eq. 27.19 for a positively charged particle, it holds for any kind of charge. Because of its velocity dependence, this force is very different from any of the other forces we have encountered so far. In particular, if the charge carrier is at rest *relative to the reference frame in which the magnetic field is measured*, the magnetic force is equal to zero. (In contrast, the electric force exerted on a charged particle is independent of the motion of the charged particle relative to the reference frame in which the electric field is measured.)

In the presence of both electric and magnetic fields, the **electromagnetic force** exerted on each charge carrier is

$$\vec{F}_p^{EB} = q\vec{E} + q\vec{v} \times \vec{B} = q(\vec{E} + \vec{v} \times \vec{B}). \tag{27.20}$$

Let us now examine what kind of trajectory a charged particle follows when it travels through a region of uniform magnetic and electric fields. We begin by examining the two special cases of a positive charge carrier traveling into a region of uniform magnetic field. When the carrier travels parallel to the direction of $\vec{B}$, as in **Figure 27.39a**, the magnetic force is zero because the vector product in Eq. 27.19 is zero when $\vec{v}$ is parallel to $\vec{B}$. (Alternatively, you can visualize the moving charge carrier as a current. As we have seen in Section 27.3, the magnetic force is zero when the current is parallel to the magnetic field.)

When the charge carrier moves perpendicular to a uniform magnetic field, as in Figure 27.39*b*, the magnetic force is nonzero. Because the magnetic force is always perpendicular to both $\vec{B}$ and $\vec{v}$, it lies in the plane of the drawing and is perpendicular to $\vec{v}$. Because the force acting on the change carrier always remains perpendicular to the direction of motion, the speed of the charge carrier does not change and we have the condition for circular motion at constant speed. The magnetic force, always directed toward the center of the circular path in which the charge carrier moves, provides the centripetal acceleration.

The equation of motion for the charge carrier is

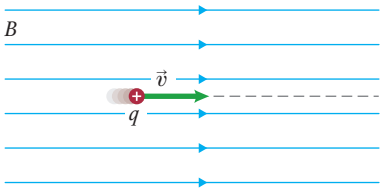
$$\Sigma\vec{F} = m\vec{a}. \tag{27.21}$$

If we ignore the force of gravity exerted on the charge carrier, the only force exerted on it is the magnetic force. Because $\vec{B}$ and $\vec{v}$ are perpendicular, the magnitude of this force is $F_p^B = |q|vB$ (Eq. 27.18). The magnitude of the charge carrier's centripetal acceleration is given by Eq. 11.15, $a_c = v^2/R$, where R is the radius of its circular trajectory. Therefore

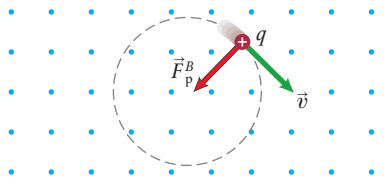
$$|q|vB = \frac{mv^2}{R}. \tag{27.22}$$

Figure 27.39 A charged particle moving in a uniform magnetic field travels (*a*) in a straight line when its velocity is parallel to the field and (*b*) in a circle when the two are perpendicular.

(*a*) Particle's velocity is parallel to field



(*b*) Particle's velocity is perpendicular to field



$\vec{B}$ out of page

Solving for R, we obtain for the radius of the trajectory

$$R = \frac{mv}{|q|B}. \tag{27.23}$$

Because the ratio $m/|q|$ is fixed for a given charge carrier, we see that the radius of its trajectory depends on the charge carrier's speed. In a given magnetic field, fast charge carriers move in larger circles than slow ones of the same type. Interestingly, the time interval needed for one full revolution is independent of the carrier's speed. This time interval is equal to the circumference of the trajectory divided by the carrier's speed:

$$T = \frac{2\pi R}{v} = \frac{2\pi}{v}\frac{mv}{|q|B} = \frac{2\pi m}{|q|B}. \tag{27.24}$$

The corresponding angular frequency is

$$\omega = 2\pi f = \frac{2\pi}{T} = \frac{|q|B}{m}. \tag{27.25}$$

This angular frequency is sometimes called the *cyclotron frequency* after a type of particle accelerator, called a *cyclotron*, in which particles are accelerated between successive semicircular trajectories.

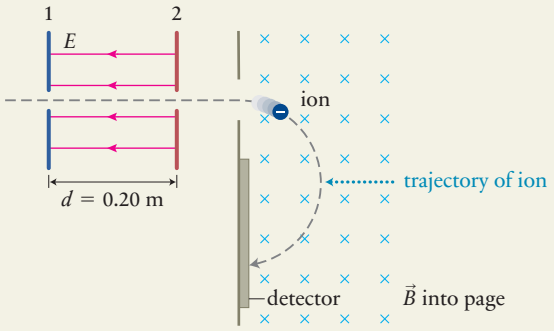


27.17 A proton and an electron travel through a region of uniform magnetic field $\vec{B}$. If their speeds are the same, what is the ratio R_p/R_e of the radii of their circular paths through the field?

EXAMPLE 27.4 Mass spectrometer

Figure 27.40 shows a schematic of a device, called a *mass spectrometer*, for determining the mass of ions or other charged particles. The ions that enter the mass spectrometer are first accelerated by an electric field and then deflected by a magnetic field. The mass of the ions is obtained from the position at which they hit a detector after being deflected by the magnetic field.

Figure 27.40 Example 27.4.



In a certain mass spectrometer, the electric field is caused by a potential difference of 10 kV across a distance $d = 0.20$ m between plates 1 and 2, and the magnitude of the magnetic field is $B = 0.20$ T. An oxygen ion with a charge q of $-2e$ ($1e = 1.6022 \times 10^{-19}$ C) and a mass m of 16 atomic mass units (1 atomic mass unit = 1 u = 1.6605×10^{-27} kg) enters the electric field with negligible initial velocity. At what distance from the point of entry into the magnetic field does the ion hit the detector?

❶ **GETTING STARTED** From the schematic I see that the ion's motion consists of two parts: a linear motion at constant acceleration in the uniform electric field, followed by circular motion at constant speed in the uniform magnetic field. Because the drawing shows the ion entering the magnetic field perpendicular to the plane of the detector, the ion traces out a half circle before hitting the detector. The distance we must determine is therefore equal to the diameter of the circular trajectory of the ion.

❷ **DEVISE PLAN** Equation 27.23 gives the radius of the circular trajectory. I am given B, m, and q, but I don't know the speed v at which the ion enters the magnetic field. This speed is equal to the final speed the ion acquires after accelerating in the electric field. I know from Eq. 25.17 how much electrostatic work is done on the ion as it traverses the potential difference between plates 1 and 2 that set up the electric field. According to Eq. 25.17, I have $W_{ion}(1 \to 2) = -qV_{12}$. This work increases the kinetic energy of the ion. Because the initial kinetic energy is zero, I have $W_{ion}(1 \to 2) = \Delta K = K_f$. I know q, V_{12}, and m, so I can now calculate v from $K_f = \frac{1}{2}mv^2$ and then use that value in Eq. 27.23 to obtain the answer to this question.

❸ **EXECUTE PLAN** The electrostatic work done on the ion is

$$W_{ion}(1 \to 2) = -qV_{12} = 2eV_{12}.$$

From the expression for kinetic energy, I obtain for the ion's speed as it enters the magnetic field

$$v = \sqrt{\frac{K_f}{\frac{1}{2}m}} = \sqrt{\frac{4eV_{12}}{m}}.$$

QUANTITATIVE TOOLS

Substituting this speed and the magnitude of the charge q on the oxygen ion into Eq. 27.23, I obtain for the diameter of the ion's circular trajectory

$$2R = \frac{2mv}{|q|B} = \frac{2m}{2eB}\sqrt{\frac{4eV_{12}}{m}} = \frac{2}{B}\sqrt{\frac{mV_{12}}{e}}$$

$$= \frac{2}{0.20\ \text{T}}\sqrt{\frac{16(1.66\times10^{-27}\ \text{kg})(1.0\times10^{4}\ \text{V})}{1.60\times10^{-19}\ \text{C}}}$$

$$= 0.41\ \text{m}.\ \checkmark$$

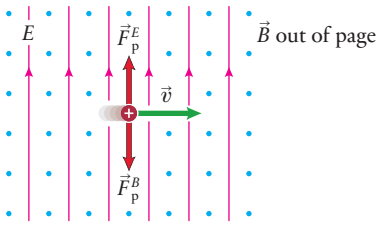
❹ **EVALUATE RESULT** To be measurable in a laboratory, the distance from the point of entry at which the ions hit the detector must be neither too small nor too great. Given these constraints, the distance I obtained appears reasonable.

Let us next consider the trajectory of a charged particle in combined uniform electric and magnetic fields. If the charged particle has a velocity component perpendicular to the magnetic field, it is subject to a magnetic force in addition to an electric force. With an appropriate choice of fields, it is possible for the two forces to add up to zero, leaving the charged particle undisturbed. This arrangement is illustrated in **Figure 27.41**: A positive charged particle moves at right angles to both $\vec{E}$ and $\vec{B}$, which are perpendicular to each other. If the two forces are of equal magnitude, we have $F_{\text{p}}^{B} = F_{\text{p}}^{E}$, and so from Eqs. 27.18 and 23.6 $|q|vB = |q|E$ or

$$v = \frac{E}{B} \quad \text{(electric and magnetic force cancel)}. \tag{27.26}$$

Thus, if we adjust the magnitudes of the electric and magnetic fields so that one charged particle goes through undeflected, then any other charged particle moving at the same velocity also passes through the fields undeflected, *regardless of its mass or charge*. Consequently, the setup shown in Figure 27.41 serves as a *velocity selector*: If charge carriers traveling at different speeds are injected from the left into a region where an electric field and a magnetic field are perpendicular to each other, only those whose speed satisfies Eq. 27.26 make it through without being deviated. Another important application of the cancellation of electric and magnetic forces is discussed in the box "The Hall effect" on page 962.

Figure 27.41 Charged particles whose speed satisfies Eq. 27.26 move in a straight line through magnetic and electric fields that are oriented perpendicular to each other.

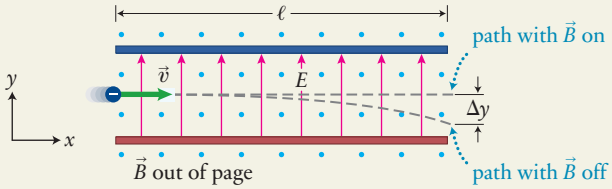


27.18 In Figure 27.41, do $\vec{F}_{\text{p}}^{B}$ and $\vec{F}_{\text{p}}^{E}$ still cancel if the charged particle (*a*) carries a negative charge, (*b*) travels in the opposite direction, (*c*) travels at a slight angle to the two fields?

EXAMPLE 27.5 The mass of the electron

Figure 27.42 shows schematically part of the apparatus used in 1897 by J. J. Thomson to determine the charge-to-mass ratio of the electron. A beam of electrons, all moving at the same speed v, enters a region of electric and/or magnetic fields. When an electric field of magnitude 1.0 kV/m and a magnetic field of magnitude 1.2×10^{-4} T are turned on, the electrons go through the device undeflected. When the magnetic field is turned off, the electrons are deflected by 3.2 mm in the negative y direction after traveling the length $\ell = 0.050$ m of the apparatus. Given that the charge of the electron is $-e = -1.60\times10^{-19}$ C, what is the mass of each electron?

Figure 27.42 Example 27.5.



❶ **GETTING STARTED** When the magnetic field is turned on, the electrons travel in a straight line, which means that

(Continued)

QUANTITATIVE TOOLS

the forces exerted by the electric and magnetic fields on the electron cancel and the electrons travel at constant speed in the $+x$ direction. When the magnetic field is turned off, the electrons undergo a constant acceleration in the $-y$ direction due to the electric field while continuing to travel at constant speed in the $+x$ direction. I know from the expression $\vec{a} = (q/m)\vec{E}$ we derived in Section 23.4 that the acceleration of the electrons in the electric field depends on their charge-to-mass ratio. If I can determine this ratio from the electrons' trajectory, I can obtain their mass.

❷ **DEVISE PLAN** Because I know the magnitudes of the electric and magnetic fields that yield a straight trajectory, I can obtain the speed v of the electrons as they enter the apparatus from Eq. 27.26. Because the electrons move in the $+x$ direction, I know that the x component of their velocity is given by $v_x = +v$ (regardless of whether the magnetic field is on or off). When the magnetic field is off, their acceleration in the y direction is given by $a_y = F_y^E/m_e = -eE/m_e$. I can solve this expression for m_e, but I don't know a_y. I do know, however, that I can use kinematics to determine a_y.

❸ **EXECUTE PLAN** Because the electron's initial velocity in the y direction is zero, the amount of deflection is given by $\Delta y = \frac{1}{2}a_y\Delta t^2$, where Δt is the time interval during which the electrons travel in the electric field. This time interval is equal to $\Delta t = \ell/v_x = \ell/v$, so that

$$a_y = \frac{2\Delta y}{(\Delta t)^2} = \frac{2\Delta y}{(\ell/v)^2} = \frac{2\Delta y\, v^2}{\ell^2} = \frac{2\Delta y(E/B)^2}{\ell^2}, \qquad (1)$$

where I have substituted Eq. 27.26 for the speed v of the electrons. Substituting $a_y = -eE/m_e$ into Eq. 1 and solving for m_e, I get

$$m_e = \frac{-eE}{a_y} = \frac{-e\ell^2B^2}{2\Delta yE}$$

$$= \frac{(-1.60\times10^{-19}\text{ C})(0.050\text{ m})^2(1.2\times10^{-4}\text{ T})^2}{2(-3.2\times10^{-3}\text{ m})(1.0\times10^{3}\text{ V/m})}$$

$$= 9.0\times10^{-31}\text{ kg.}\ \checkmark$$

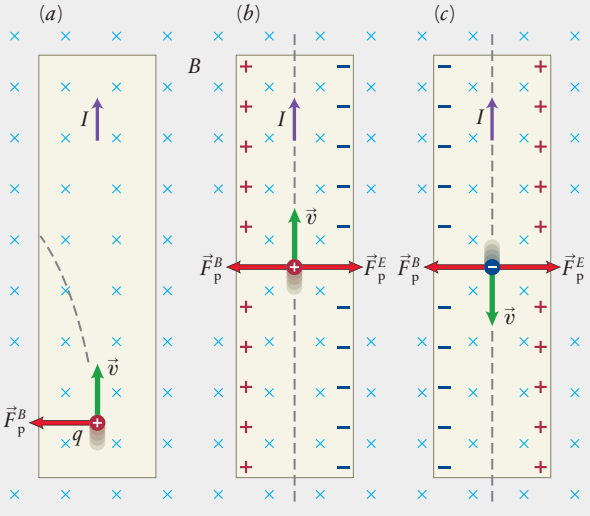
❹ **EVALUATE RESULT** The value I obtain is close to the published value of the electron mass ($m_e = 9.10938370\times10^{-31}$ kg), giving me confidence in my calculation.

The Hall effect

The canceling of electric and magnetic forces exerted on a charge carrier makes it possible to determine whether the mobile charge carriers in a conductor are positively or negatively charged. This determination is done using a phenomenon called the *Hall effect*.

Consider the rectangular conducting strip carrying an upward current illustrated in **Figure 27.43** (that is, the bottom of the strip is at a higher potential than the top). The strip is placed in a magnetic field directed into the page. If the current is caused by positive charge carriers (Figure 27.43*a*), the carriers move upward and so, according to the right-hand rule, the magnetic force acting on them is to the left. This force deflects the carriers toward the left side of the strip, where they pile up. This accumulation of positive charge carriers on the left causes an electric field to the right across the strip that exerts on the carriers an electric force that pulls them to the right. As more carriers accumulate on the left, the magnitude of this electric force increases until it becomes equal in magnitude to the magnetic force. Once the two forces are equal in magnitude, the carriers are no longer deflected and so no more carriers accumulate on the left. Any subsequent charge carriers travel in a straight line.

Figure 27.43



The accumulation of charge carriers can be determined by measuring the potential difference between the left (L) and right (R) sides of the strip. If, as illustrated in Figure 27.43*b*, positive carriers accumulate on the left, this potential difference is positive:

$$V_{RL} = V_L - V_R > 0.$$

As illustrated in Figure 27.43*c*, however, if the mobile charge carriers causing the current are negatively charged, V_{RL} is *negative*. This is so because an upward current means that negative charge carriers move downward and, as you can verify using the right-hand rule, the magnetic force exerted on these carriers is also to the left. Thus, regardless of

QUANTITATIVE TOOLS

the sign of their charge, the mobile charge carriers pile up on the left. If they are positively charged, V_{RL} is positive; if they are negatively charged, V_{RL} is negative. Experiments on strips of common metals always yield a negative V_{RL}, showing that the mobile charge carriers in these metals are negatively charged, as we stated earlier (see Section 22.3 and Checkpoint 22.15).

27.19 (*a*) Express the magnitude of the electric field inside the strip in Figure 27.43 in terms of the width w of the strip and the potential difference V_{RL}. (*b*) Given the magnitude B of the magnetic field, what is the speed at which the charge carriers travel? (*c*) Show how this information, together with Eq. 27.16, can be used to determine the number density n of the charge carriers.

27.8 Magnetism and electricity unified

In this section we quantitatively examine the situation discussed in the latter part of Section 27.4—the interaction between a current-carrying wire and a positively charged particle moving parallel to the wire. For simplicity, we again let the charged particle move at the same speed v as the electrons in the wire.* In the Earth reference frame in Figure 27.27 the wire is electrically neutral: The linear charge density of the fixed positively charged ions $\lambda_{proper} > 0$ is equal in magnitude to that of the electrons, $-\lambda_{proper}$. The interaction between the wire and the charged particle is thus purely magnetic.

As we saw in Section 27.4, the wire appears to be positively charged to observer M, who moves along with the electrons. According to M, the average distance between the positively charged ions has decreased by the factor

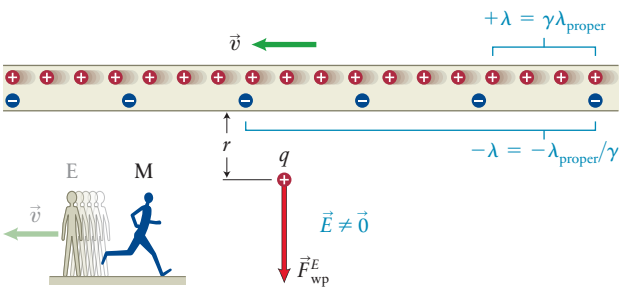
$$\gamma = \frac{1}{\sqrt{1 - v^2/c_0^2}} \tag{27.27}$$

due to length contraction (Eqs. 14.28 and 14.6; $\gamma > 1$). Thus the charge density of the ions is increased and has a magnitude $\lambda_{M\ ions} = \gamma\lambda_{proper}$ (**Figure 27.44*a***). The electrons, on the other hand, are at rest relative to M and moving relative to E. Therefore the average distance between the electrons according to E must be decreased by the factor γ in Eq. 27.27 relative to the distance observed by M. Because to observer E in the Earth reference frame the wire is electrically neutral, the charge density of the

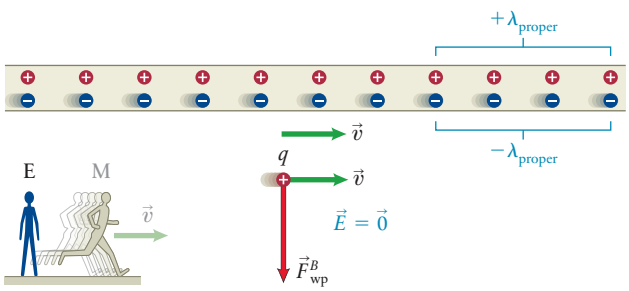
*In the more general case where the charged particle does not move at the same speed as the electrons in the wire, the algebra becomes more complicated but the conclusions remain the same.

Figure 27.44 A positively charged particle moving parallel to a current-carrying wire, as seen by two observers in motion relative to each other.

(*a*) **Reference frame of electrons in wire (observer M)**



(*b*) **Earth reference frame (observer E)**



electrons in that reference frame must be the negative of that of the ions: $-\lambda_{\text{proper}}$. According to M, therefore, the charge density must be $\lambda_{\text{M electrons}} = -\lambda_{\text{proper}}/\gamma$. Consequently, the combined charge density of the ions and the electrons in the moving reference frame of observer M is

$$\lambda_{\text{M}} = \lambda_{\text{M ions}} + \lambda_{\text{M electrons}} = \gamma\lambda_{\text{proper}} - \frac{\lambda_{\text{proper}}}{\gamma} = \lambda_{\text{proper}}\left(\gamma - \frac{1}{\gamma}\right). \quad (27.28)$$

Because $\gamma > 1$, the term in parentheses is positive, making the combined charge density positive. Using Eq. 14.9 we can rewrite the term in parentheses as

$$\gamma - \frac{1}{\gamma} = \gamma\left(1 - \frac{1}{\gamma^2}\right) = \gamma\,\frac{v^2}{c_0^2}. \quad (27.29)$$

Substituting Eq. 27.29 into Eq. 27.28, we see that according to M the wire has a (nonzero) charge density equal to

$$\lambda_{\text{M}} = \lambda_{\text{proper}}\gamma\,\frac{v^2}{c_0^2}. \quad (27.30)$$

Using the expression $E = 2k\lambda/r$ for the magnitude of the electric field of a charged wire (see Exercise 24.7), we obtain for the magnitude of the electric field at a distance r from the wire according to M

$$E_{\text{M}} = \frac{2k\lambda_{\text{M}}}{r} = \frac{2k\lambda_{\text{proper}}\gamma v^2}{rc_0^2}. \quad (27.31)$$

(Because the distance to the wire is perpendicular to the direction of motion, both observers measure the same distance r and so we need no subscript on r.)

According to observer M, the electric force exerted by the wire on the charged particle therefore has a magnitude

$$F^E_{\text{Mwp}} = |q|E_{\text{M}} = |q|\,\frac{2k\lambda_{\text{proper}}\gamma v^2}{rc_0^2} \quad (27.32)$$

and points perpendicular to and away from the wire.

Now we return to the Earth reference frame, as illustrated in Figure 27.44*b*. In this reference frame, the combined charge density of the ions and electrons is

$$\lambda_{\text{ions}} + \lambda_{\text{electrons}} = \lambda_{\text{proper}} - \lambda_{\text{proper}} = 0, \quad (27.33)$$

and so the electric field outside the wire is zero. Consequently, the force between the wire and the charged particle cannot be due to an electric interaction. We know experimentally, however, that according to observer E the current-carrying wire exerts a *magnetic* force on the particle. The direction of this force is also perpendicular to and away from the wire. Is this magnetic force the same as the electric force seen by M?

To answer this question we first need to determine the relationship between the forces measured by two observers in motion relative to each other. To determine this relationship, we shall use the definition of force given in Eq. 8.4, $\Sigma\vec{F} \equiv d\vec{p}/dt$, where $\Sigma\vec{F} = \vec{F}_{\text{wp}}$. The vector $d\vec{p}$ points in the same direction as the force—perpendicular to the relative velocity of the two frames—and so must be the same for both observers: $d\vec{p} = d\vec{p}_{\text{M}}$. Because of time dilation, however, we have from Eq. 14.13, $dt = \gamma dt_{\text{M}}$,

where dt is the infinitesimal time interval measured by observer E and dt_M is the corresponding time interval measured by observer M (because M moves along with the particle, this is a proper time interval). The magnitude of the force exerted by the wire on the particle measured by observer E is thus

$$F_{wp} = \frac{dp}{dt} = \frac{dp_M}{\gamma dt_M} = \frac{F_{Mwp}}{\gamma}, \tag{27.34}$$

where F_{Mwp} is the magnitude of the force exerted by the wire on the particle measured by observer M. To observer M this force is electric in nature and its magnitude is given by Eq. 27.32. Substituting Eq. 27.32 on the right in Eq. 27.34 thus yields

$$F_{wp} = |q| \frac{2k\lambda_{proper} v^2}{rc_0^2}. \tag{27.35}$$

The quantity $\lambda_{proper} v$ is just the current I because if the electrons advance by a distance d in a time interval Δt, a quantity of charge $|q| = \lambda_{proper} d$ flows through the wire in that time interval. So the rate at which the charge carriers flow is

$$I \equiv \left| \frac{q}{\Delta t} \right| = \lambda_{proper} \frac{d}{\Delta t} = \lambda_{proper} v. \tag{27.36}$$

This result means that we can write Eq. 27.35 as

$$F_{wp} = |q| v \frac{2kI}{rc_0^2}. \tag{27.37}$$

Equation 27.37 gives the magnitude of the force exerted by the current-carrying wire on the particle measured by observer E. The magnitude depends on the current I through the wire, on the speed v and charge q of the particle, and on the distance r between the wire and the particle. Observer E interprets this force as a magnetic force due to the magnetic field caused by the current through the wire. From Eq. 27.18 we see that the magnitude of the magnetic force exerted by the wire on the moving particle is $F^B_{wp} = |q|vB$ (because $\vec{v}$ and $\vec{B}$ are at right angles, $\sin\theta = 1$). Substituting this expression into Eq. 27.37 and dividing both sides by qv, we see that the magnitude of the magnetic field according to observer E is

$$B = \frac{2kI}{rc_0^2} = \frac{2k}{c_0^2} \frac{I}{r}. \tag{27.38}$$

In words, the magnitude B of the magnetic field is proportional to the current I and inversely proportional to the distance r to the wire, with the proportionality constant being $2k/c_0^2$. As we shall see in the next chapter, experiments confirm this dependence. Here we see that Coulomb's law, together with special relativity, *requires* a magnetic field of this form. In other words, electricity and magnetism are two aspects of the same interaction, not two different interactions.



27.20 Explain why the $1/r$ dependence expressed in Eq. 27.38 is consistent with the symmetry of the wire causing the magnetic field.

Chapter Summary

Magnetic field (Sections 27.1, 27.2)

Concepts Every **magnet** has two *magnetic poles*. The pole attracted to Earth's North Pole is called a *north pole;* the magnet's opposite pole is called a *south pole*.

Like magnetic poles repel each other, and opposite magnetic poles attract each other. Earth's North Pole is a magnetic south pole because it attracts the north pole of a compass needle.

Magnetic materials are attracted to both types of magnetic poles. The presence of a magnet induces **magnetic polarization** in magnetic materials.

A magnet is surrounded by a **magnetic field.** This field exerts a force on the poles of other magnets and is represented by **magnetic field lines.** At any position, the direction of the magnetic field is tangent to the magnetic field line passing through that position. Magnetic field lines form loops that exit a magnet at its north pole and reenter the magnet at its south pole. Near (but outside of) a magnet, magnetic field lines point away from north poles and toward south poles. At any location, the magnitude of the magnetic field is proportional to the field line density at that location.

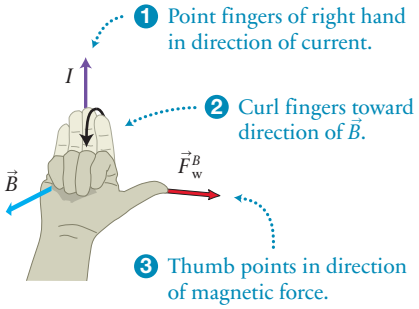
Currents and magnetic fields (Sections 27.3, 27.5)

Concepts A flow of charged particles, called a **current,** causes a magnetic field.

The *direction of current* through a conductor is the direction in which positive charge carriers would flow, which is from high potential to low potential.

Right-hand current rule: If you point the thumb of your right hand in the direction of the current, the fingers curl in the direction of the magnetic field produced by that current.

Right-hand rule for vector product $I\vec{\ell} \times \vec{B}$



Quantitative Tools If a particle carrying charge dq passes a given section of a conductor in a time interval dt, the current I through that section is

$$I \equiv \frac{dq}{dt}. \tag{27.2}$$

We measure current in **amperes** (A), where

$$1\ \text{C} \equiv 1\ \text{A}\cdot\text{s}. \tag{27.3}$$

If a straight wire of length ℓ carrying a current I perpendicular to a uniform magnetic field $\vec{B}$ experiences a force $\vec{F}^B_{w,\,max}$, the magnitude of the magnetic field is

$$B \equiv \frac{F^B_{w,\,max}}{|I|\,\ell}. \tag{27.5}$$

The SI unit of the magnetic field is the **tesla** (T), where

$$1\ \text{T} \equiv 1\ \text{N}/(\text{A}\cdot\text{m}) = 1\ \text{kg}/(\text{s}^2\cdot\text{A}). \tag{27.6}$$

The **magnetic force** $\vec{F}^B_w$ exerted on a straight wire in a uniform magnetic field is

$$\vec{F}^B_w = I\vec{\ell} \times \vec{B}, \tag{27.8}$$

where the direction of $\vec{\ell}$ is the direction of the current through the wire. The direction of $\vec{F}^B_w$ is obtained by curling the fingers of the right hand from the direction of the current to $\vec{B}$.

Table 27.1 Right-hand rules in magnetism

	Thumb points in direction of	Fingers curl
Current rule	Current	Along magnetic field
Force rule	Magnetic force	From current to magnetic field

It is sometimes helpful to write out Eq. 27.8 component by component:

$$F^B_{wx} = I\ell_y B_z - I\ell_z B_y$$
$$F^B_{wy} = I\ell_z B_x - I\ell_x B_z$$
$$F^B_{wz} = I\ell_x B_y - I\ell_y B_x.$$

Magnetic flux (Section 27.6)

Concepts The **magnetic flux** through a surface is a measure of the number of magnetic field lines crossing the surface. The magnetic flux through any closed surface is zero because magnetic field lines always form closed loops.

Quantitative Tools The **magnetic flux** Φ_B through a surface is

$$\Phi_B \equiv \int \vec{B} \cdot d\vec{A}, \tag{27.10}$$

where the integral is taken over the surface.

Magnetic flux is measured in **webers** (Wb), where

$$1\ \text{Wb} \equiv 1\ \text{T} \cdot \text{m}^2 = 1\ \text{m}^2 \cdot \text{kg}/(\text{s}^2 \cdot \text{A}).$$

For any closed surface, **Gauss's law for magnetism** is

$$\Phi_B = \oint \vec{B} \cdot d\vec{A} = 0. \tag{27.11}$$

Forces exerted on moving particles (Section 27.7)

Quantitative Tools If a charged particle of charge q moves with velocity $\vec{v}$ in the presence of an electric field and a magnetic field, the **electromagnetic force** $\vec{F}^{EB}_p$ exerted on the particle by the fields is

$$\vec{F}^{EB}_p = q(\vec{E} + \vec{v} \times \vec{B}). \tag{27.20}$$

It is sometimes helpful to write out Eq. 27.20 component by component:

$$F^{EB}_{px} = q(E_x + v_y B_z - v_z B_y), \quad F^{EB}_{py} = q(E_y + v_z B_x - v_x B_z),$$
$$F^{EB}_{pz} = q(E_z + v_x B_y - v_y B_x).$$

A charged particle of charge q and mass m moving at speed v perpendicular to a uniform magnetic field $\vec{B}$ follows a circular path of radius R given by

$$R = \frac{mv}{|q|B}. \tag{27.23}$$

The time interval T for one revolution is

$$T = \frac{2\pi m}{|q|B}, \tag{27.24}$$

and the angular frequency ω is

$$\omega = \frac{|q|B}{m}. \tag{27.25}$$

Magnetism and electricity unified (Sections 27.4, 27.8)

Concepts The interaction observed between charge carriers depends on their motion relative to the observer: The interaction can be purely electric, purely magnetic, or a combination of the two. Magnetic and electric interactions are two aspects of one *electromagnetic* interaction, with magnetism being a relativistic correction to the electric interaction.

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

27.1 Magnetism

1. Which way does the north pole of a compass needle point in the Southern Hemisphere? •
2. You place two magnets side by side but, instead of attracting, the magnets repel each other. Provide a possible explanation of your observation. •
3. How are the elementary magnets aligned in the magnet in Figure P27.3? •

Figure P27.3

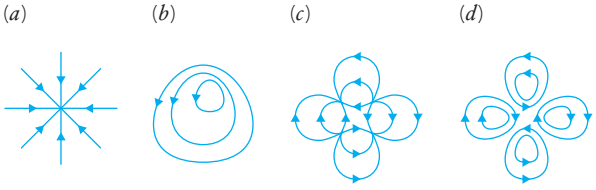


4. Is it possible to separate the north pole from the south pole of a bar magnet by cutting it in half? Explain your reasoning. •
5. You are given three bars of metal. Two of them are magnets, and the third is made of a magnetic material but is not magnetized. Describe how, using only the three bars, you could determine which of them is not a magnet. ••
6. You are given two metal rods. One is a magnet, and the other is made of magnetic material but does not have the elementary magnets aligned. Using no other objects, how can you determine which is the magnet? ••
7. Sketch and describe the magnetic poles of a spherical piece of uniformly magnetized material. •••

27.2 Magnetic fields

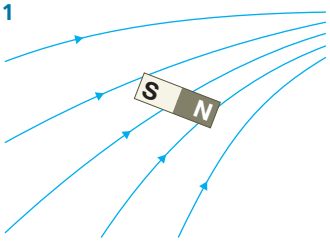
8. Which field line patterns in Figure P27.8 can represent a magnetic field? •

Figure P27.8



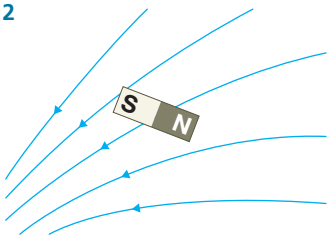
9. A bar magnet is enclosed by a spherical surface. (*a*) What is the magnetic flux through the entire surface? (*b*) If the magnetic flux through the hemisphere closest to the magnet's north pole is positive, is the flux through the hemisphere closest to the magnet's south pole positive, negative, or zero? (*c*) The magnet is now moved so that its north half is outside the surface. What is the flux through the entire surface now? (*d*) The portion of the magnet outside the surface is sawed off. What is the flux through the surface now? •
10. Is the field line pattern created by a magnetic dipole the same as the field line pattern created by an electric dipole? Draw both field line patterns. ••
11. Describe what happens to a bar magnet placed in the nonuniform external magnetic field shown in Figure P27.11. ••

Figure P27.11



12. Describe what happens to a bar magnet placed in the nonuniform external magnetic field shown in Figure P27.12. ••

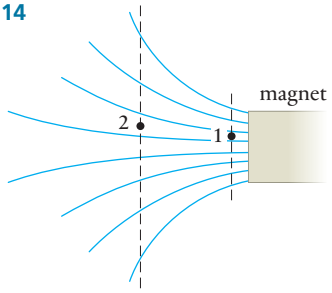
Figure P27.12



13. How can the direction of the magnetic field at a certain point along a curved field be determined? •

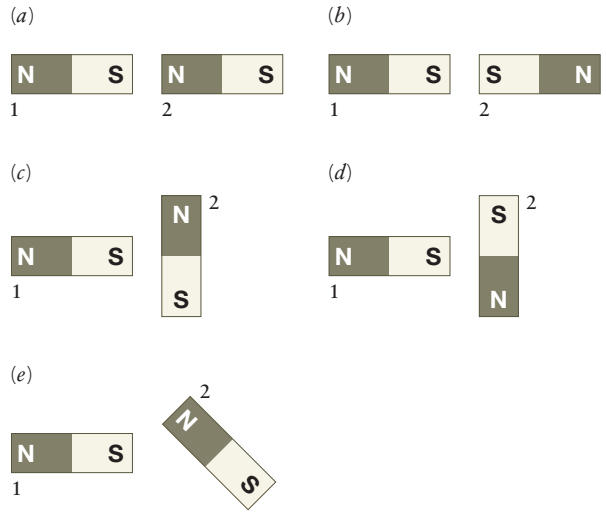
14. Estimate the magnitude of the magnetic field at location 2 in Figure P27.14 if the magnetic field magnitude at location 1 is 0.27 T. Note that real magnetic field lines spread out in three dimensions. The spreading is shown in only two dimensions here, but you must imagine that the spreading also occurs into and out of the page. ●●

Figure P27.14



15. Rank the magnet pairs in Figure P27.15 in order of the magnitude of torque on magnet 2, smallest torque first. Assume all the magnets are equal in strength and the spacing between magnets is comparable in all pairs. ●●●

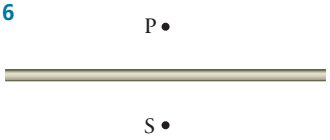
Figure P27.15



27.3 Charge flow and magnetism

16. The long, straight current-carrying wire of Figure P27.16 lies in the plane of the page, and the magnetic field it produces at position P points out of the page. (*a*) In what direction does the magnetic field point at position S? (*b*) What is the current direction in the wire? ●

Figure P27.16



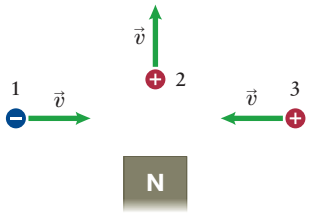
17. Figure P27.17 shows a long, straight current-carrying wire running perpendicular to the plane of the page. The current produces a magnetic field that points to the right at position P. (*a*) In what direction does the magnetic field point at position S? (*b*) What is the current direction in the wire? ●

Figure P27.17



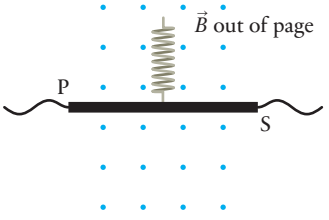
18. Figure P27.18 shows three particles passing near the north end of a bar magnet. Particle 1 is an electron, and particles 2 and 3 are protons. All three particles move at the same speed. (*a*) Determine the direction of the magnetic force exerted on each particle as it passes over the magnet. (*b*) What is the direction of the force (if any) exerted by each particle on the magnet? ●

Figure P27.18



19. Figure P27.19 shows a conducting rod suspended from a spring in a region where a uniform magnetic field points horizontally out of the page. The rod can be supplied with current by two thin wires connecting its ends to the terminals of a battery (not shown). (*a*) Does the tension in the spring depend on whether or not the wires are connected to the battery? (*b*) Does the tension depend on which wire is connected to which battery terminal? If so, which wire should be connected to the positive terminal in order to increase the tension in the spring? ●●

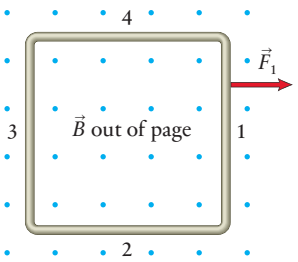
Figure P27.19



20. A wire is coiled in the shape of a helical spring with closely spaced turns. (*a*) When current is passed through it, does the coil tend to lengthen, shorten, or stay the same length? (*b*) Does your answer depend on the current direction? ●●

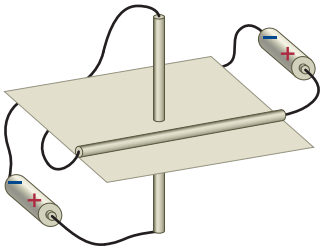
21. The square loop of wire in Figure P27.21 carries a current, and an external magnetic field is directed out of the page everywhere. If the magnetic force exerted on side 1 is to the right, determine (*a*) the direction of the current in the loop and (*b*) the directions of the magnetic forces exerted on sides 2, 3, and 4. ●●

Figure P27.21



22. The two metal rods in Figure P27.22 are perpendicular to each other. Describe the magnetic force each rod exerts on the other, and then describe the torque caused by each force. ●●●

Figure P27.22



27.4 Magnetism and relativity

23. If two charged particles M and S are at rest relative to each other, there is no magnetic force between them. Suppose instead that particle M is moving relative to particle S while S is at rest in the Earth reference frame. Is a magnetic force exerted on S? ●

24. A positively charged particle is at rest on the positive z axis in reference frame S. Reference frame S′ is moving along the positive x axis of S, reference frame S″ is moving along the negative x axis, and reference frame S‴ is moving along the y axis, all with speed v relative to S. The axes of all the reference frames are oriented the same way, and their origins coincide at instant $t = 0$, when observers in each reference frame make measurements. In which reference frames do observers measure an electric field? ●

25. Write expressions for the magnitudes and directions of the electric fields measured in Problem 24. ●●

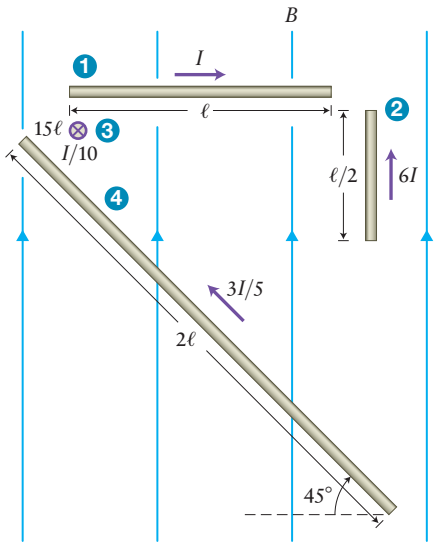
26. Compare the magnitudes and directions of the magnetic fields measured in Problem 24. ●●

27.5 Current and magnetism

27. A 0.40-m wire carries a current of 1.7 A and is at an angle to a 0.20-T uniform external magnetic field. What must be the angle between the wire and the magnetic field so that the field exerts a force of 0.068 N on the wire? ●●

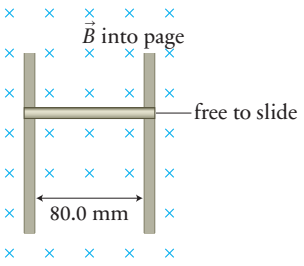
28. In Figure P27.28, an external magnetic field is directed upward throughout a region that contains four current-carrying wires having the lengths and currents shown. Rank the wires according to the magnitude of the magnetic force exerted on them by the external field, smallest force first. Ignore any interactions between the wires. ●

Figure P27.28



29. Two vertical parallel rails made of material that is an electrical conductor are 80.0 mm apart (Figure P27.29). A wire 80.0 mm long is free to slide along the rails, which allow a current to be maintained through the wire as it slides. An external magnetic field of magnitude 0.250 T is directed into the page in Figure P27.29. If the mass of the wire is 4.00 g, what are the current and current direction in the wire if the wire doesn't fall under gravity? ●●

Figure P27.29



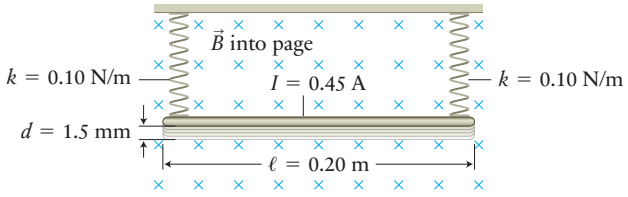
30. A current-carrying wire is bent into a circular loop of radius R and lies in an xy plane. A uniform external magnetic field in the $+z$ direction exists throughout the plane of the loop. What is the magnetic force exerted by the external field on the loop? ●●

31. Draw an xyz coordinate system with the x axis pointing horizontally to the right, the y axis pointing up the page, and the z axis pointing out of the page. Show a current-carrying wire lying on the x axis, and a 2.5-A current through the wire directed in the positive x direction. (*a*) If the charge carriers are only electrons, how many of them pass through a cross-sectional area of the wire each second? (*b*) In which direction are the electrons moving? (*c*) If this wire is in a 0.20-T external magnetic field directed into the page, what is the force magnitude per unit length exerted by the external field on the wire? (*d*) On your diagram, show the direction of the magnetic field and the direction of the force exerted by the external field on the wire. ●●

32. A square loop of side $a = 50$ cm is suspended from the left arm of a scale, which is kept at equilibrium by adding weights to the right side. The lower half of the loop is in a uniform magnetic field of magnitude 0.4 T, orthogonal to the surface of the loop, but the upper half is outside the magnetic field. (*a*) It is observed that when a current I flows in the loop, one needs to add 25 g to the right arm to bring the scale back to equilibrium. Find the current flowing in the loop. (*b*) If the direction of current were now to be reversed, what would need to be done to bring the scale back to equilibrium? ●●

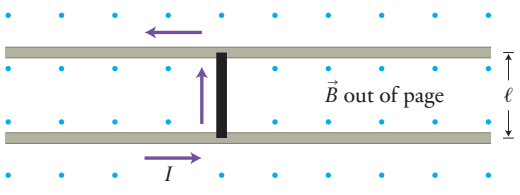
33. Figure P27.33 shows the arrangement we looked at in Example 27.2: a metal bar 0.20 m long suspended from two springs, each having a spring constant $k = 0.10$ N/m. Initially there is no current through the bar, and it is suspended at rest below the ceiling. With a current of 0.45 A, the bar rises a distance $d = 1.5$ mm. Suppose that the current is turned off, the bar drops down to its original position, and we attach a 5.0-mg piece of plastic to the center of the bar. Now how much current must there be through the bar to make it rise a distance $d = 1.5$ mm? ●●

Figure P27.33



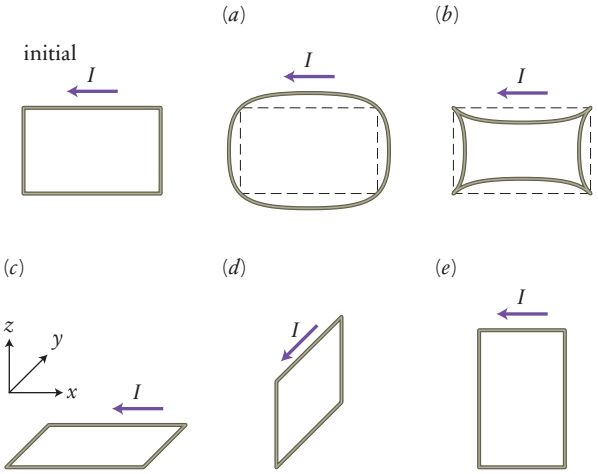
34. Two horizontal parallel conducting rods are connected such that a conducting crossbar free to slide along them has a constant current I running through it (Figure P27.34). The rods are separated by a distance ℓ and are in an external uniform magnetic field of magnitude B directed out of the page. The crossbar has a length ℓ and mass m. (*a*) In which direction does the crossbar move? (*b*) If there is a coefficient of static friction μ_s between rods and crossbar, what is the minimum current I_0 necessary for the crossbar to move? ●●

Figure P27.34



35. The top left portion of Figure P27.35 shows a current-carrying wire shaped into a rectangular loop that is very flexible. The loop is mounted on a base (not shown) that allows it to spin in any direction, and in this way any forces that cause the loop to move, spin, or distort can be detected. If the current direction through the loop is counterclockwise as indicated, what magnetic field direction(s) could cause the shape and orientation changes shown in parts (*a*)–(*e*) of Figure P27.35? If no magnetic field is necessary to make a shown change, write "none." If there is no way a magnetic field could cause a shown change, write "not possible." ●●

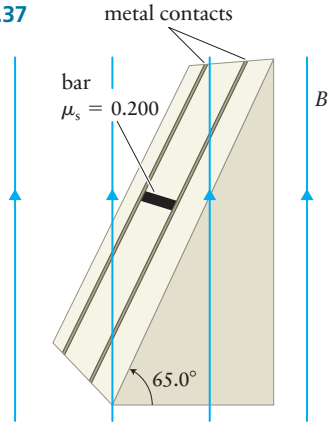
Figure P27.35



36. An arbitrary-shaped tangle of wire is connected such that it carries a current I_0 from position $\vec{r}_1$ to position $\vec{r}_2$ in a region where there is an external uniform magnetic field $\vec{B}_0$. Show that the magnetic force exerted by the external field on the wire is given by $\vec{F}_w^B = I_0(\vec{r}_2 - \vec{r}_1) \times \vec{B}_0$. ●●●

37. A 1.00-m metal bar that has a mass of 0.900 kg is initially pinned in place on an incline 65.0° above the horizontal (Figure P27.37). There is a 0.850-T magnetic field directed upward in the region around the bar, and metal contacts along the incline allow a current through the bar. If the coefficient of static friction between bar and contacts is 0.200, (*a*) what are the minimum magnitude and the direction of the current that must pass through the bar to keep it from moving once it is released? (*b*) Is there a maximum current that can be used in that same direction that will also hold the bar in place? ●●●

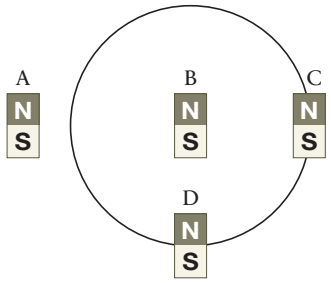
Figure P27.37



27.6 Magnetic flux

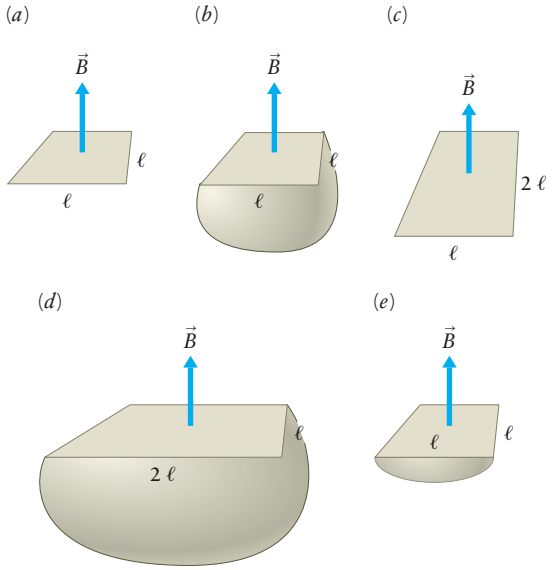
38. Figure P27.38 shows a bar magnet placed at four positions on and near a spherical shell. Rank the positions according to the amount of magnetic flux through the shell, smallest flux first. •

Figure P27.38



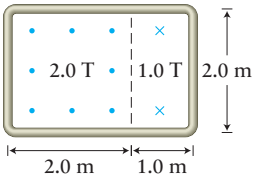
39. Figure P27.39 shows five objects, all placed in the same uniform, upward-directed external magnetic field. Rank the objects according to the amount of magnetic flux through them, smallest flux first. •

Figure P27.39



40. A square loop of side length 1 m is placed on a wooden table in a uniform magnetic field of magnitude 0.5 T. The greatest magnetic flux through the loop is measured when the loop is flat on the table. (*a*) What is the direction of the magnetic field? (*b*) What is the flux through the loop when it is tilted such that the plane of the loop makes an angle of 30° with the table? ••
41. A circular loop of diameter 150 mm is placed on a wooden table that makes angle of 24.5° with a uniform magnetic field. What must be the magnitude of the magnetic field if the magnetic flux through the loop is $4.20 \times 10^{-4}\ \text{T} \cdot \text{m}^2$? ••
42. A square loop of wire has a perimeter of 60 cm and is oriented such that two of its parallel sides form a 30° angle with the horizontal. A uniform horizontal magnetic field of magnitude 0.2 T passes through the loop. (*a*) Calculate the magnetic flux through the loop. (*b*) Another loop lies in the same plane but has an irregular shape resembling that of a starfish. Its area equals that of the square loop. Calculate the flux through this loop. ••
43. The loop of Figure P27.43 is partially in a region where there is a 2.0-T magnetic field pointing out of the page and partially in a region where there is a 1.0-T magnetic field pointing into the page. Calculate the magnetic flux through the loop. ••

Figure P27.43

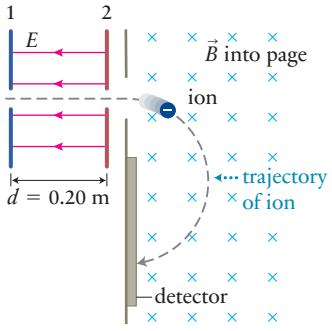


44. A hemispherical bowl of radius R is placed in a uniform magnetic field that has magnitude B_0 and is in the positive z direction. The open top of the bowl is in the xy plane. Obtain an expression for the magnetic flux through the hemispherical surface of the bowl. ••

27.7 Moving particles in electric and magnetic fields

45. Can a charged particle move in a region of uniform magnetic field without experiencing an external force? If yes, how? If no, why not? •
46. Figure P27.46 shows the semicircular path through which charged particles travel in the magnetic field of a

Figure P27.46



mass spectrometer. If the particles are oxygen ions carrying a charge of $-2e$ and the magnetic field magnitude is 0.20 T, how long does it take each ion to travel the semicircular path through the magnetic field? The mass of an O^{2-} ion is 2.6×10^{-26} kg. •

47. An electron is moving at speed 7.0×10^{10} m/s perpendicular to a uniform magnetic field of magnitude 0.80 T. (*a*) What is the diameter of the circular path that the electron follows? Calculate the electron's (*b*) angular frequency and (*c*) period of motion. •••

48. An alpha particle ($m = 6.64 \times 10^{-27}$ kg), which has twice the charge and approximately four times the mass of a proton, moves in a circular orbit 100 mm in diameter that is perpendicular to a uniform 0.75-T magnetic field. Calculate the alpha particle's (*a*) angular frequency and period of motion, (*b*) speed, and (*c*) kinetic energy. •••

49. An electron that has been accelerated through a 100-V potential difference moves in a circular orbit perpendicular to a uniform magnetic field of magnitude 0.4 T. For this electron, determine (*a*) its speed, (*b*) its angular frequency and period of motion, and (*c*) the radius of its orbit. ••

50. A deuteron is a charged particle that has the same charge as a proton but approximately twice the mass. An alpha particle has twice the charge and approximately four times the mass of a proton. For this problem, assume the mass ratios are exactly 2 and 4. If a proton, a deuteron, and an alpha particle move in a uniform magnetic field in circles having the same radii, compare their (*a*) speeds, (*b*) kinetic energies, and (*c*) angular momentum magnitudes. ••

51. The isotopes magnesium-24 (mass 3.983×10^{-26} kg) and magnesium-26 (mass 4.315×10^{-26} kg) are to be separated using a mass spectrometer in which the magnetic field magnitude is 0.4 T. The ions used are the ones that have lost one electron each, $^{24}Mg^{+}$ and $^{26}Mg^{+}$. The resolution of your detector screen is 3 mm. What is the minimum value of the potential difference through which these ions must be accelerated to see them distinctly on the detector screen? ••

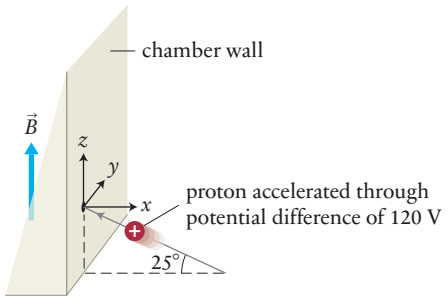
52. A particle that has mass m and charge q enters a uniform magnetic field of magnitude B. The initial velocity v of the particle forms an angle of 30° with the magnetic field. After some time t, the velocity of the particle has the same magnitude and direction that it had initially. What is the distance d traveled by the particle over this time period? ••

53. A horizontal metal strip 1.0 mm thick and 10 mm wide carries a 12-A current along its length, and both the length and the width are perpendicular to a uniform magnetic field of magnitude 1.0 T that is directed vertically up. The potential difference across the strip width is 3 μV. (*a*) If the side of the strip width that is to your left as you look in the direction of the current is at the higher potential, what is the sign of the mobile charge carriers? What is (*b*) the speed and (*c*) the number density of the charge carriers? ••

54. The cross section of a copper strip is 1.0 mm thick and 20 mm wide. There is a 10-A current through this cross section, with the charge carriers traveling down the length of the strip. The strip is placed in a uniform magnetic field that has a magnitude of 2.0 T and is directed perpendicular to both the length and the width of the strip. If the number density of free electrons in copper is 8.47×10^{19} mm^{-3}, calculate (*a*) the speed of the electrons in the strip and (*b*) the potential difference across the strip width. ••

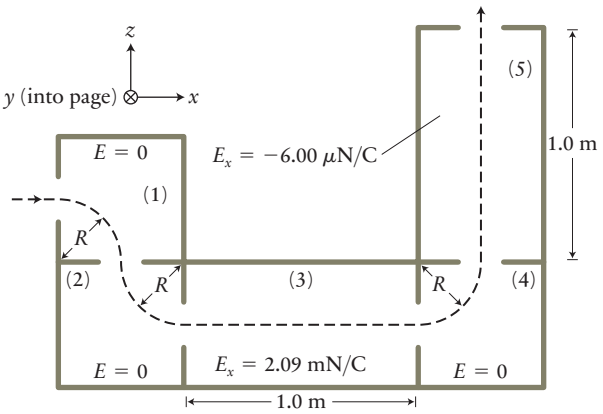
55. A proton is accelerated through a potential difference of 120 V, as shown in Figure P27.55, and fired into a chamber. There is no electric field in the chamber, but there is a 0.15-T magnetic field in the positive z direction. The proton enters the chamber at a 25° angle above the horizontal, and this entry location is the origin of the xyz coordinate system. Determine the coordinates of the location at which the proton will later strike this same wall from the inside. ••

Figure P27.55



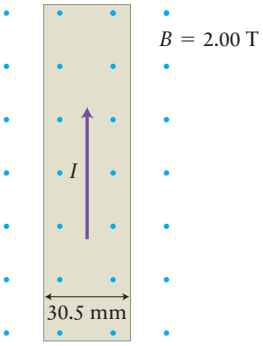
56. A beam of protons enters the network of five chambers shown in Figure P27.56 with an initial speed of 300 m/s and moves through the network along the path indicated by the dashed black line. In each chamber, the electric field is as specified, and the radius of curvature R is 0.40 m wherever it is shown. For each chamber, calculate the component of the magnetic field perpendicular to the path of the protons. •••

Figure P27.56



57. Electrons are made to flow through the copper strip of Figure P27.57. The strip's cross section is 1.00 mm high and 30.5 mm wide, and the strip is placed in a 2.00-T magnetic field that is directed out of the page. If the number of mobile electrons per unit volume in the copper is 8.46×10^{28} m^{-3} and the current in the strip is 12 A, calculate the potential difference across the width of the strip when conditions have been allowed to equilibrate. •••

Figure P27.57



27.8 Magnetism and electricity unified

58. On average, the number density of free electrons in copper is 8.46×10^{19} mm^{-3}. (*a*) Calculate what the linear charge density λ for a copper wire 1.00 mm in radius would be if this quantity of electrons were missing. (*b*) Assume this copper wire is electrically neutral in the Earth reference frame, in which it is at rest and carrying a current. Calculate the linear charge density λ' of the copper wire in a reference frame moving along with the electrons if the electrons are moving at 4.70×10^{-4} m/s. ••

59. A square loop of wire 1.00 m on a side is at rest and has a linear charge density of zero in the Earth reference frame. The loop lies in the *xy* plane of an *xyz* coordinate system, and there is a clockwise electric current through the loop as seen by a person looking down the positive *z* axis to the *xy* plane. The electrons move at speed *v*, and an observer M, moving at this same speed *v*, approaches the loop along the *y* axis (moving in the positive *y* direction). In which side(s) of the loop ($+x$, $+y$, $-x$, $-y$) does M see the linear charge density as being due (*a*) to surplus positive charge carriers and (*b*) to surplus negative charge carriers? ••

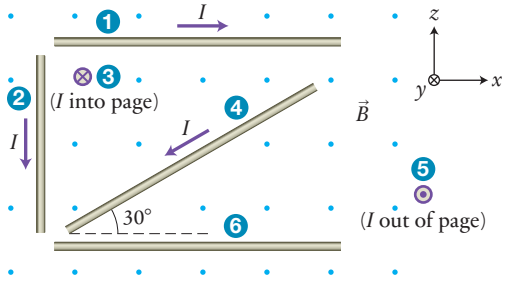
60. Suppose electrons move through a copper wire at speed *v*. Call the linear charge densities in the Earth reference frame λ_{Ep} for the fixed positive ions in the wire and λ_{En} for the (negative) electrons. Observer E in the Earth reference frame (which is also the reference frame of the positive ions) measures the wire to be electrically neutral. Observer M is moving along with the electrons (in the same direction as the electrons and at speed *v*). Calculate the linear charge density of the wire as measured by M in terms of λ_{En} if the speed is (*a*) 3.0×10^5 m/s and (*b*) 3.0 mm/s (an achievable speed for electrons in copper wire). ••

61. A long wire that is at rest in the Earth reference frame initially carries no current. Observer E in this reference frame measures linear charge densities of $\lambda_{En} = 45.00\ \mu$C/m for the (negative) electrons in the wire and $\lambda_{Ep} = 60.00\ \mu$C/m for the positive ions. A current is created as electrons are made to move through the wire at speed *v*, and observer M moves along with the electrons at their speed *v*. What value must *v* have in order for M to measure an electric field of magnitude $E = 3.00 \times 10^7$ N/C at a distance 0.0100 m above the wire? •••

Additional Problems

62. In Figure P27.62, an external magnetic field is directed out of the page, and six wires are placed, one at a time, in this field. Each wire carries a current in the direction indicated; if there is no indicated direction, the current is zero. Determine the direction of the magnetic force exerted by the magnetic field on each wire. •

Figure P27.62



63. Three charged particles initially have identical masses, charges, and speeds and are traveling perpendicular to the same magnetic field. Because they are all in the same magnetic field and have identical masses and charges, all particles have the same initial cyclotron frequency and revolve with the same period *T*. The speed of particle 1 is doubled; particle 2 enters a region in which the magnetic field has been reduced by a factor of 2; and particle 3 collides with and sticks to a neutral particle of the same mass, causing a change in speed. Determine the new period of each particle in terms of the initial period *T*. •

64. At the equator, the direction of Earth's magnetic field is horizontal and to the north, and the magnitude is 3.5×10^{-5} T. (*a*) What is the magnetic flux through a circular loop of radius 0.10 m, lying flat on the ground at the equator? The loop is now balanced on its edge and the axis that runs perpendicular to the plane of the loop is pointed in different directions. What is the magnetic flux through the loop when this axis is pointed (*b*) northwest, (*c*) due north, and (*d*) west? •

65. When a wire 2.0 m long carries a 10-A current in the $+x$ direction in a uniform external magnetic field, the magnetic force exerted by the field on the wire is 3 N

in the $+y$ direction. When the wire is rotated until the charge carriers travel in the $+y$ direction, the magnitude of the magnetic force exerted by the field on the wire is 5 N. What are the magnitude and direction of the magnetic field? ••

66. Blood contains both positive and negative ions. In a certain patient being tested, the speed of these ions in a large artery is measured at 0.60 m/s. The patient is placed in a uniform magnetic field of magnitude 0.20 T. If a potential difference of 1.0×10^{-3} V is measured across the diameter of the artery, what is that diameter? ••

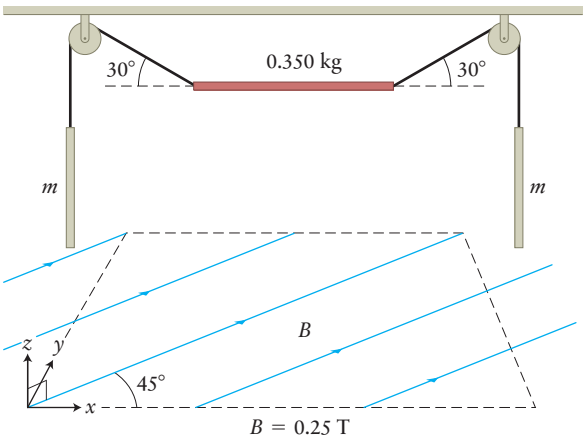
67. Electrons enter a region of perpendicular electric and magnetic fields (similar to Example 27.5). The region occupied by the fields is 40.0 mm long in the x direction, where x is taken to be the original direction of motion. After passing through this region, the electrons then travel an additional 300 mm in the x direction before hitting a phosphorescent screen that lights up where the electrons strike it. The electrons pass undeflected through the region of perpendicular fields when $E = 2.0 \times 10^3$ N/C and $B = 1.2 \times 10^{-5}$ T. (*a*) Determine the speed of the undeflected electrons. (*b*) If the magnetic field is turned off but the electric field is left on, by what distance have the electrons been deflected when they reach the phosphorescent screen? ••

68. A positively charged particle of mass m and charge $+q$ moves in a circle of radius r in a uniform magnetic field of magnitude B. How does the radius of the particle's motion change when (*a*) the magnitude of the magnetic field is reduced into half, (*b*) the mass of the charge is doubled, and (*c*) the sign of the charge is changed to negative? ••

69. You are designing magnetic motors. A colleague insists that like magnetic poles attract—north poles attract north poles and south poles attract south poles. You feel the need to remove this idea from his head. You possess three unlabeled bar magnets. How can you use them to convince him? ••• CR

70. Figure P27.70 shows a 1.0-m, 0.350-kg horizontal metal rod attached to two ropes, each of which makes an angle of 30° with the horizontal. Each rope then drapes over a pulley and is attached to a vertical rod of mass m. Assume the pulleys have negligible mass. There is a 0.25-T magnetic field in the region of the rod, parallel to the xy plane, oriented 45° from the long axis of the horizontal rod. When the horizontal rod is made to carry a current I, suddenly the rod moves upward until the ropes holding it make an angle of 5.0° with the horizontal. What was the value of I? •••

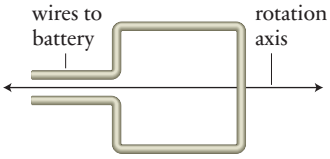
Figure P27.70



71. A wire bent into a semicircle of radius R lies in a plane that is perpendicular to a uniform external magnetic field $\vec{B}$. If the wire carries a current I, what are the magnitude and direction of the magnetic force exerted by the external field on the wire? •••

72. You are working with the square wire loop shown in Figure P27.72. The loop lies in a magnetic field, is connected to a battery, and is attached to an axis that runs through its center and is parallel to the plane of the loop. Initially the loop is held in place, but once it is released, it immediately rotates 180° about this axis. You know that magnetic fields can induce a torque on a current-carrying wire, but you are surprised by the rotation of exactly 180°. In an attempt to explain why this happened, you determine the direction of the current in the loop and the direction of the magnetic field in the region of the loop. Are multiple answers possible? Is there any way to keep the loop rotating in the same direction? ••• CR

Figure P27.72





28

Magnetic Fields of Charged Particles in Motion

CONCEPTS

28.1 Source of the magnetic field
28.2 Current loops and spin magnetism
28.3 Magnetic dipole moment and torque
28.4 Ampèrian paths

QUANTITATIVE TOOLS

28.5 Ampère's law
28.6 Solenoids and toroids
28.7 Magnetic fields due to currents
28.8 Magnetic field of a moving charged particle

In this chapter we investigate further the relationship between the motion of charged particles and the occurrence of magnetic fields. As we shall see, *all* magnetism is due to charged particles in motion—whether moving along a straight line or spinning about an axis. It takes a moving or spinning charged particle to create a magnetic field, and it a takes another moving or spinning charged particle to "feel" that magnetic field. We shall also discuss various methods for creating magnetic fields, which have wide-ranging applications in electromechanical machines and instruments.

28.1 Source of the magnetic field

As we saw in Chapter 27, magnetic interactions take place between magnets, current-carrying wires, and moving charged particles. **Figure 28.1** summarizes the interactions we have encountered so far. Figures 28.1*a–c* show the interactions between magnets and current-carrying wires. The sideways interaction between a magnet and a current-carrying wire (Figure 28.1*b*) is unlike any other interaction we have encountered. The forces between the wire and the magnet are not central; they do not point directly from one object to the other. As we saw in Section 27.7, the magnetic force exerted on a current-carrying wire is the sum of the magnetic forces exerted on many individual moving charge carriers. Similarly the magnetic field due to a current-carrying wire is the sum of the magnetic fields of many individual moving charge carriers. Figures 28.1*d* and *e* illustrate the magnetic interactions of moving charged particles. Note that for two charged particles moving parallel to each other (Figure 28.1*e*), there is, in addition to an attractive magnetic force, a (much greater) repulsive electric force.

It is important to note that the magnetic interaction depends on the state of motion of the charged particles. No magnetic interaction occurs between a bar magnet and a stationary charged particle (Figure 28.1*f*) or between two stationary charged particles (Figure 28.1*g*). These observations suggest that the motion of charged particles might be the origin of *all* magnetism. There are two problems with this assumption, however. First, the magnetic field of a wire carrying a constant current looks very different from that of a bar magnet. (Compare Figures 27.13 and 27.19.) Second, there is no obvious motion of charged particles in a piece of magnetic material.

Figure 28.2*a* shows the magnetic field line pattern of a straight wire carrying a constant current. The lines form circles centered on the wire, circles that reflect the cylindrical symmetry of the wire (the symmetry of an infinite cylinder, see Section 24.4). The horizontal distance between adjacent circles is smaller near the wire, where the magnitude of the field is greater.

A single moving charged particle does not have cylindrical symmetry because, unlike for an infinitely long straight wire, moving the particle up or down along its line of motion changes the physical situation. Because of the *circular* symmetry of the situation, the field still forms

Figure 28.1 Summary of magnetic interactions. Notice that stationary charged particles do not engage in magnetic interactions.

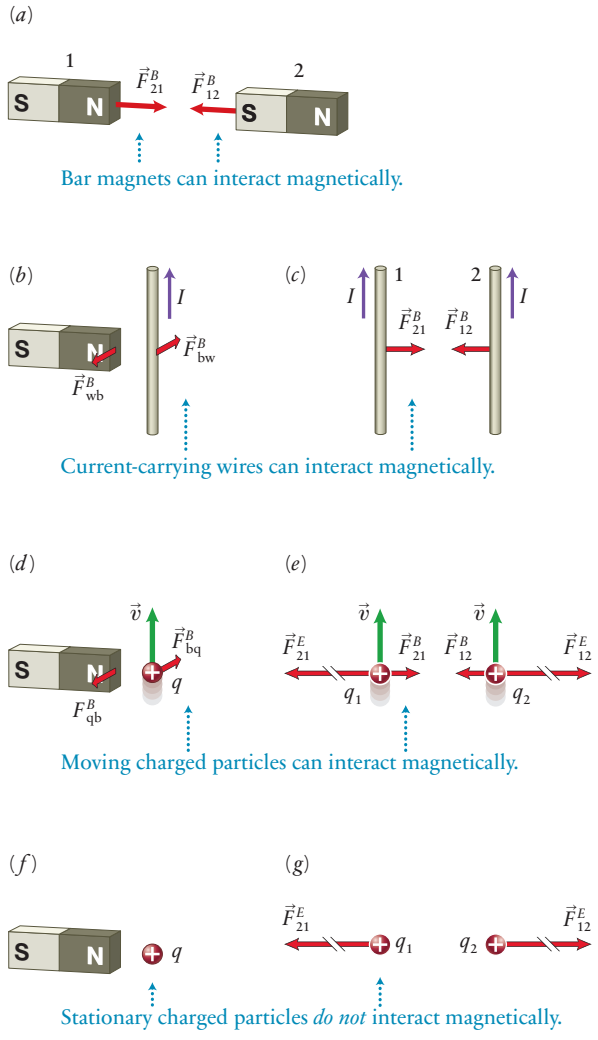
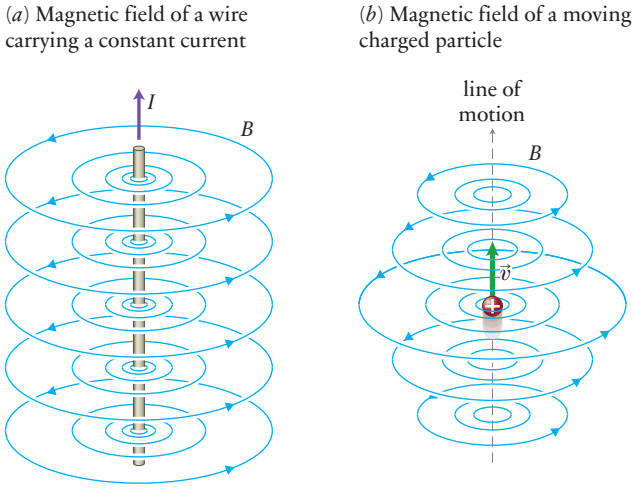


Figure 28.2 Comparing the magnetic fields of a current-carrying wire and a moving charged particle.

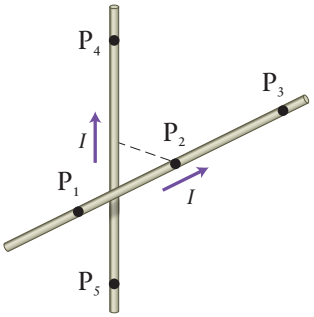


CONCEPTS

circles around the line of motion, but the magnitude of the field at a fixed distance from the particle's line of motion decreases if one moves away from the particle (Figure 28.2*b*). It is not at all obvious how the field pattern shown in Figure 28.2*b* could give rise to the magnetic field of a bar magnet; there are certainly no poles in the magnetic field of the moving charged particle.

28.1 Make a sketch showing the directions of the magnetic forces exerted on each other by (*a*) an electron moving in the same direction as the current through a wire, (*b*) a moving charged particle and a stationary charged particle, and (*c*) two current-carrying wires at a right angle to each other as illustrated in **Figure 28.3**. (Hint: Determine the forces exerted at points P_1 through P_5.)

Figure 28.3 Checkpoint 28.1*c*.



28.2 Current loops and spin magnetism

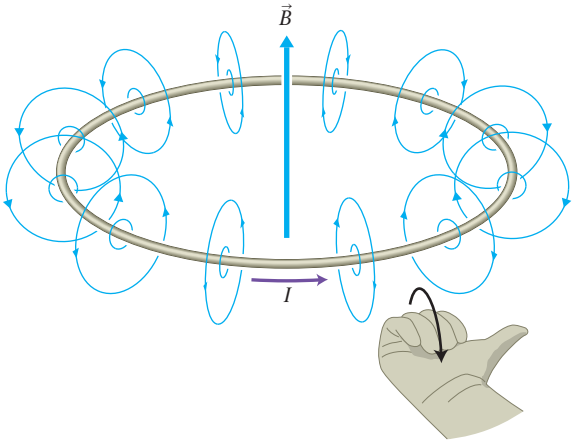
The circular pattern of magnetic field lines around a wire carrying a constant current suggests a method for generating a strong magnetic field: If a wire carrying a constant current is bent into a loop as shown in **Figure 28.4**, all the magnetic field lines inside the loop point in the same direction, reinforcing one another.*

What does the magnetic field of such a current-carrying loop, called a **current loop** for short, look like? To answer this question, we treat the current loop as a collection of small segments of a current-carrying wire and determine the direction of the magnetic field at various points around the loop. As we did in Chapter 27, we shall assume all currents to be constant in the remainder of this chapter.

We begin by considering the magnetic field due to a small segment of the current loop at a point on the central axis that passes perpendicularly through the face of the loop (point A in **Figure 28.5**). Segment 1 carries a current that points into the page. The magnetic field lines of this segment are concentric circles centered on the segment. Using the right-hand current rule, we see that the magnetic field curls clockwise in the plane of the drawing, and

*For now, we'll ignore how to make charge carriers flow through such a loop. In Section 28.6 we'll discuss physical arrangements that accomplish the situation illustrated in Figure 28.4.

Figure 28.4 The magnetic field of a wire loop carrying a constant current. The magnetic fields from all parts of the loop reinforce one another in the center of the loop.



so the magnetic field due to segment 1 at A points up and to the right. Figure 28.5*b* shows a magnetic field line through A due to segment 2. Because the current through segment 2 points out of the page, this field line curls counterclockwise, and so the magnetic field due to segment 2 at A points up and to the left.

Figure 28.5*c* shows the contributions from segments 1 and 2 together; because their horizontal components cancel, the vector sum of $\vec{B}_1$ and $\vec{B}_2$ points vertically up. The same arguments can be applied to any other pair of segments lying on opposite sides of the current loop. Therefore the magnetic field due to the entire current loop points vertically up at A.

Now consider point C at the center of the current loop. As illustrated in Figure 28.5*d*, the magnetic fields due to segments 1 and 2 point straight up there, too. The same is true for all other segments, and so the field at point C also points straight up. We can repeat the procedure for point D below the current loop, and as shown in Figure 28.5*e*, the magnetic field there also points up.

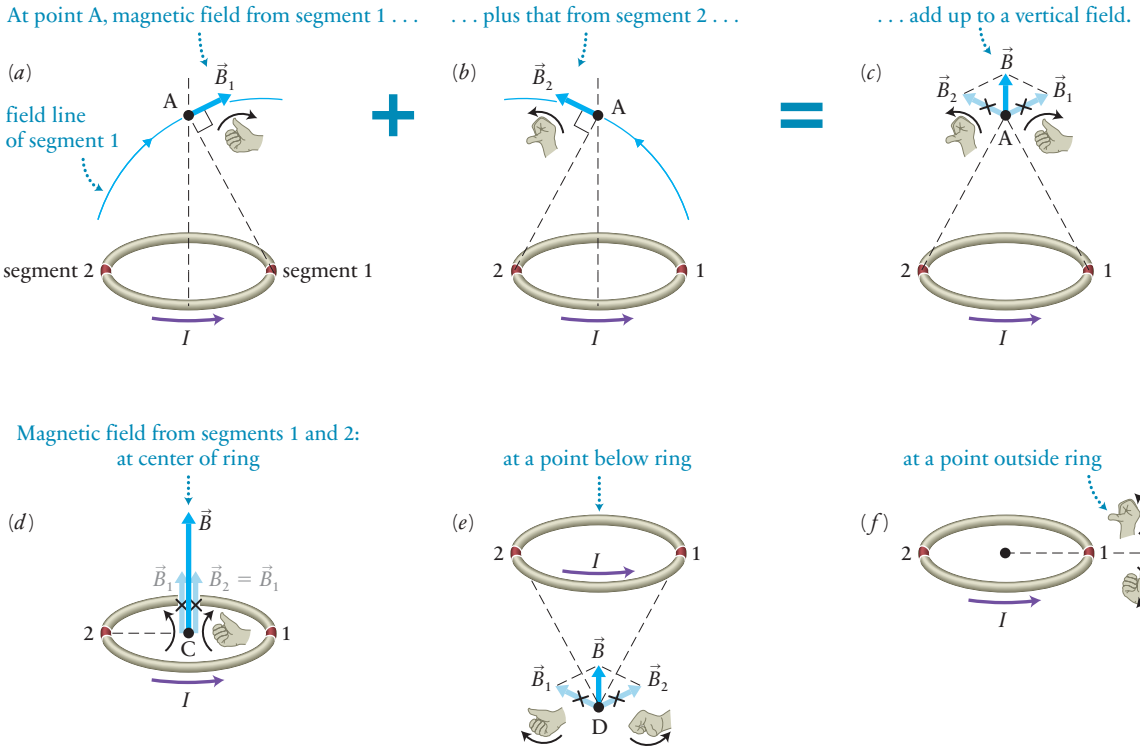
To determine the direction of the magnetic field at a point outside the current loop, consider point G in Figure 28.5*f*. Using the right-hand current rule, you can verify that the field due to segment 1 points vertically down and the field due to segment 2 points vertically up. Because G is closer to 1 than it is to 2, the magnetic field due to 1 is stronger, and so the sum of the two fields points down.

28.2 What is the direction of the magnetic field at a point vertically (*a*) above and (*b*) below segment 1 in Figure 28.5?

The complete magnetic field pattern of the current loop, obtained by determining the magnetic fields for many more points, is shown in **Figure 28.6***a*. Close to the wire, the field lines are circular, but as you move farther away from the wire, the circles get squashed inside the loop and stretched outside due to the

CONCEPTS

Figure 28.5 Mapping the magnetic field of a current loop. The magnetic field contributions from (*a*) segment 1 and (*b*) segment 2 at A (*c*) add up to a vertical field. Magnetic fields are shown at (*d*) point C at the center of the ring, (*e*) point D below the ring, and (*f*) point G to the right of the ring. Note that in all cases the magnetic field of each segment is perpendicular to the line connecting that segment to the point at which we are determining the field.



contributions of other parts of the loop to the magnetic field.

As you may have noticed, the field line pattern of a current loop resembles that of a bar magnet (Figure 28.6*b*). Indeed, if you shrink the size of both the current loop and the bar magnet, their magnetic field line patterns become identical (Figure 28.6*c*). The magnetic field line pattern in Figure 28.6*c* is that of an infinitesimally small magnetic dipole.

Given that a current loop produces a magnetic field similar to that of a bar magnet, is the magnetic field of a bar magnet then perhaps due to tiny current loops inside the magnet? More precisely, are elementary magnets (see Section 27.1) simply tiny current loops?

The connection between current loops and elementary magnets becomes clearer once we realize that a current loop is nothing but an amount of charge that revolves around an axis. Consider, for example, a positively charged ring spinning around a vertical axis through its center (**Figure 28.7**). The spinning charged

Figure 28.6 The magnetic field lines of a current loop (*a*) resemble that of a dipole (*b*, *c*).

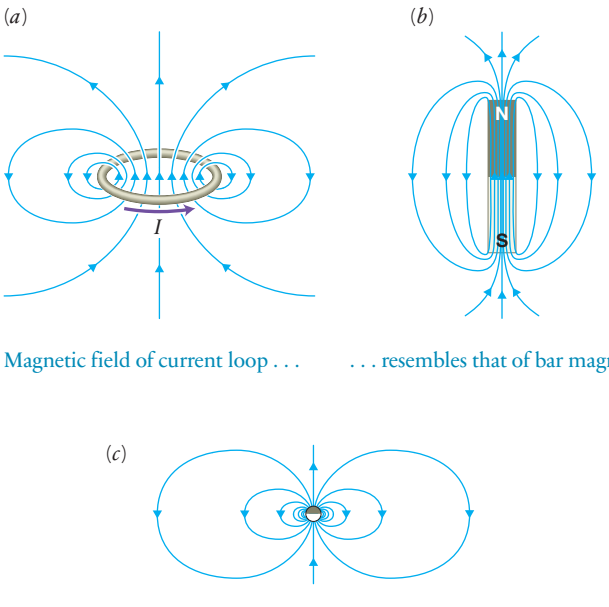
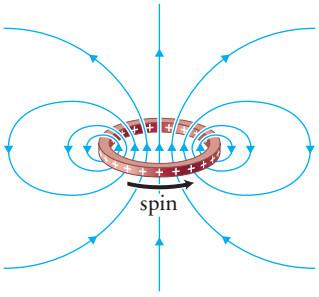


Figure 28.7 The magnetic field of a charged spinning ring is identical to that of a current loop.



CONCEPTS

particles cause a current moving in a circle exactly like the current through the circular loop in Figure 28.5. If we let the radius of the ring approach zero, the ring becomes a spinning charged particle, and its magnetic field pattern approaches that of the magnetic dipole illustrated in Figure 28.6*c*. This surprising result tells us:

A spinning charged particle produces a magnetic field identical to that produced by an infinitesimally small magnetic dipole.

Experiments show that most elementary particles, such as electrons and protons, possess an intrinsic angular momentum—as if they permanently spin around—and such spinning motion indeed would produce a magnetic field of the form shown in Figure 28.6*c*. Because of the combined intrinsic angular momentum of these elementary particles inside atoms, certain atoms have a magnetic field, causing them to be the elementary magnets we discussed earlier. The reason we cannot separate north and south magnetic poles is therefore a direct consequence of the fact that the magnetic field of a particle with intrinsic angular momentum is that of an infinitesimally small magnetic dipole.

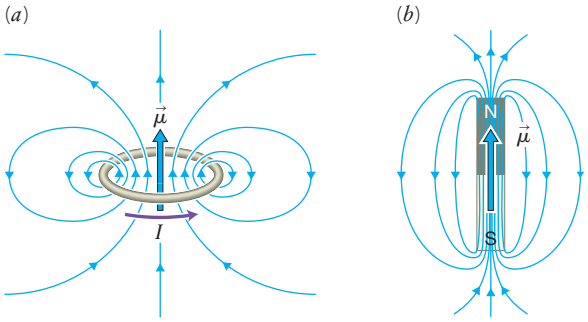


28.3 Suppose a negatively charged ring is placed directly above the positively charged ring in Figure 28.7. If both rings spin in the same direction, is the magnetic interaction between them attractive or repulsive?

28.3 Magnetic dipole moment and torque

To specify the orientation of a magnetic dipole we introduce the **magnetic dipole moment.** This vector, represented by the Greek letter $\vec{\mu}$ (mu), is defined to point, like a compass needle, along the direction of the magnetic field through the center of the dipole (**Figure 28.8**). For a bar magnet the magnetic dipole moment points from the south pole to the north pole. To determine the direction of $\vec{\mu}$ for a current loop, you can use a right-hand rule: When you curl the fingers of your right hand along the direction of the current, your thumb points in the direction of $\vec{\mu}$.

Figure 28.8 The magnetic dipole moment points in the direction of the magnetic field through the center of the dipole.



We now have three right-hand rules in magnetism, illustrated in **Figure 28.9** and summarized in **Table 28.1** together with the right-hand vector product rule. What helps distinguish these right-hand rules from one another is looking at which quantity curls: For a current-carrying wire, it is the magnetic field that curls, whereas for a current loop, it is the current that curls. The only additional thing you need to remember is that in the case of the magnetic force exerted on a current-carrying wire, the fingers are associated with the curl from the direction of the current to the direction of the magnetic field.

Table 28.1 Right-hand rules

Right-hand rule	thumb points along	fingers curl
vector product	$\vec{C} = \vec{A} \times \vec{B}$	from $\vec{A}$ to $\vec{B}$
current rule	current	along B-field
force rule	magnetic force	from current to B-field
dipole rule	$\vec{\mu}$ (parallel to $\vec{B}$)	along current loop

CONCEPTS

Figure 28.9 Right-hand rules in magnetism.

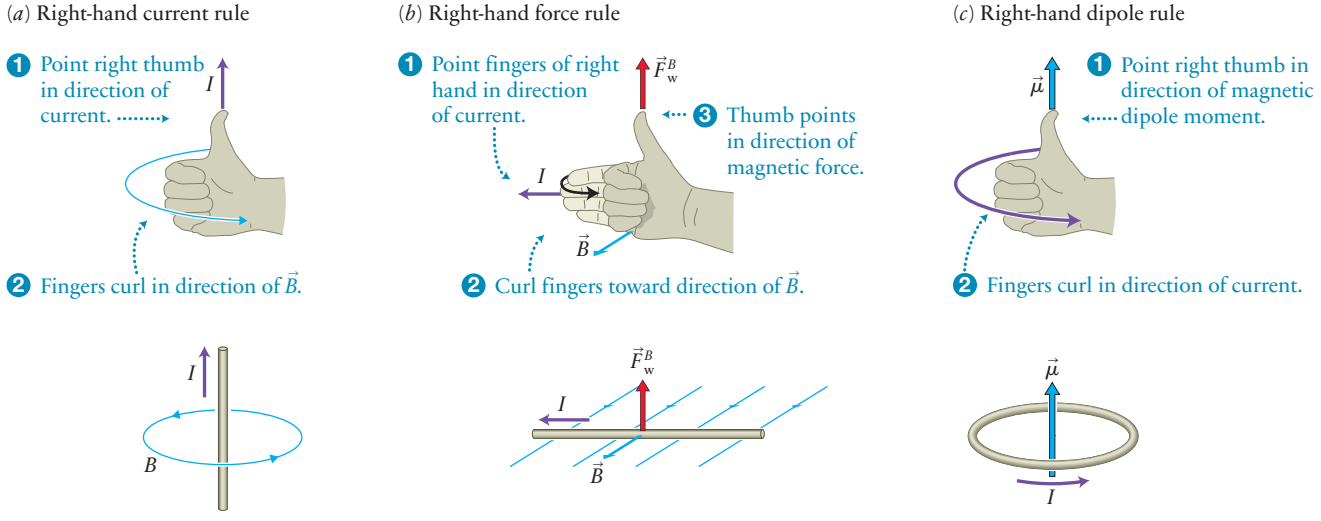


Figure 28.10 Magnetic forces exerted on a square current loop that is oriented so that its magnetic dipole moment is perpendicular to an external magnetic field.

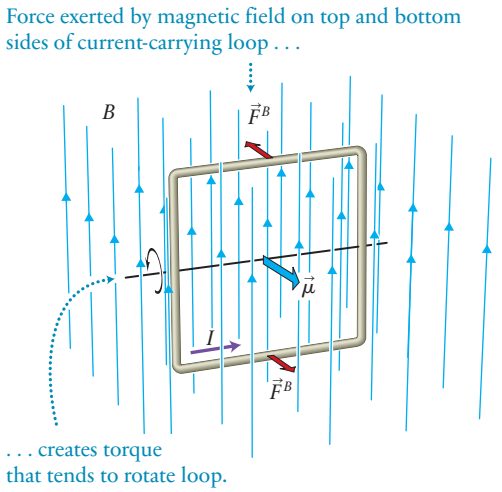
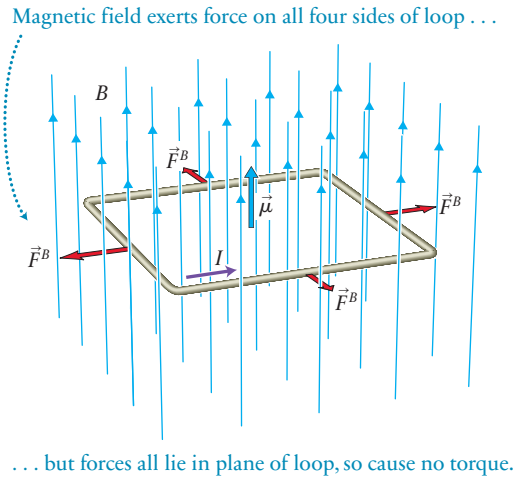


Figure 28.11 Magnetic forces exerted on a square current loop that is oriented so that its magnetic dipole moment is parallel to an external magnetic field.



Note, also, that the order of application of each rule in Figure 28.9 can be reversed to solve for a different variable. If you know the direction in which the magnetic field lines curl, for example, you can use the right-hand current rule to determine the current direction.

28.4 Does the direction of the electric field along the axis inside an electric dipole coincide with the direction of the electric dipole moment?

What happens when a current loop is placed in a magnetic field? To find out, consider a square current loop of wire placed in a uniform magnetic field with its magnetic dipole moment perpendicular to the magnetic field (**Figure 28.10**). The current loop experiences magnetic forces on the top and bottom sides but not on the vertical sides because they are parallel to the direction of the magnetic field (see Section 27.3). Using the right-hand force rule of Figure 28.9*b*, we see that the magnetic forces exerted on the top and bottom sides cause a torque that tends to rotate the loop as indicated in Figure 28.10.

When the current loop is oriented with its magnetic dipole moment parallel to the magnetic field, all four sides experience a magnetic force (**Figure 28.11**). However, because all four forces lie in the plane of the loop, none of them causes any torque. Because the magnitudes of the four forces are the same, their vector sum is zero and the loop is not accelerated sideways.

28.5 As the current loop in Figure 28.10 rotates over the first 90°, do the magnitudes of (*a*) the magnetic force exerted on the horizontal sides and (*b*) the torque caused by these forces increase, decrease, or stay the same? (*c*) As the loop rotates, do the two vertical sides experience any force, and, if so, do these forces cause any torque? (*d*) What happens to the torque as the loop rotates beyond 90°?

Summarizing the results of Checkpoint 28.5:

A current loop placed in a magnetic field tends to rotate such that the magnetic dipole moment of the loop becomes aligned with the magnetic field.

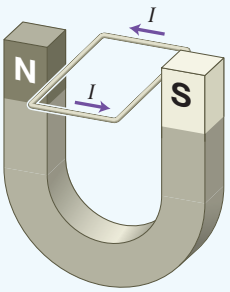
This alignment is completely analogous to the alignment of the electric dipole moment in the direction of an external electric field, which we studied in Section 23.8 (see Figure 23.32).

28.6 Suppose the square current loop in Figure 28.10 is replaced by a circular loop with a diameter equal to the width of the square loop and with the same current. Does the circular loop experience a torque? If not, why not? If so, how does this torque compare with that on the square loop?

EXAMPLE 28.1 Current loop torque

When placed between the poles of a horseshoe magnet as shown in **Figure 28.12**, does a rectangular current loop experience a torque? If so, in which direction does the loop rotate?

Figure 28.12 Example 28.1.



1 GETTING STARTED I know that a current loop placed in a magnetic field tends to rotate such that the magnetic dipole moment of the loop becomes aligned with the magnetic field. If the loop tends to rotate, it experiences a torque.

2 DEVISE PLAN The simplest way to answer this question is to look at the directions of the magnetic field $\vec{B}$ and the magnetic dipole moment $\vec{\mu}$. By definition, the magnetic field between the poles of the magnet points from the north pole to the south pole. To determine the direction of $\vec{\mu}$, I can use the right-hand dipole rule.

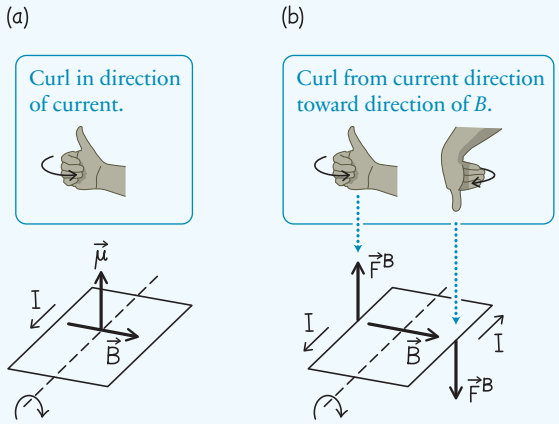
(Continued)

CONCEPTS

3 EXECUTE PLAN I begin by sketching the loop and indicating the direction of the magnetic field (**Figure 28.13*a***). To determine the direction of $\vec{\mu}$, I curl the fingers of my right hand along the direction of the current through the loop. My thumb shows that $\vec{\mu}$ points straight up. To align $\vec{\mu}$ with $\vec{B}$, therefore, the loop rotates in the direction shown by the curved arrow in Figure 28.13*a*. The loop must experience a torque in order for this rotation to occur. ✔

4 EVALUATE RESULT I can verify my answer by determining the force exerted by the magnetic field on each side of the loop and seeing if these forces cause a torque (Figure 28.13*b*). The front and rear sides experience no force because the current through them is either parallel or antiparallel to $\vec{B}$. To determine the direction of the force exerted on the left side of the loop, I point my right-hand fingers along the direction of the current through that side and curl them toward the direction of the magnetic field. My upward-pointing thumb indicates that the force exerted on the left side is upward. Applying the right-hand force rule to the right side of the loop tells me that the magnetic force exerted on that side is directed downward. The magnetic forces exerted on the left and right sides thus cause a torque that makes the loop rotate in the same direction I determined earlier.

Figure 28.13



The alignment of the magnetic dipole moment of current loops in magnetic fields is responsible for the operation of any device involving an electric motor (see the box below titled "Electric motors").

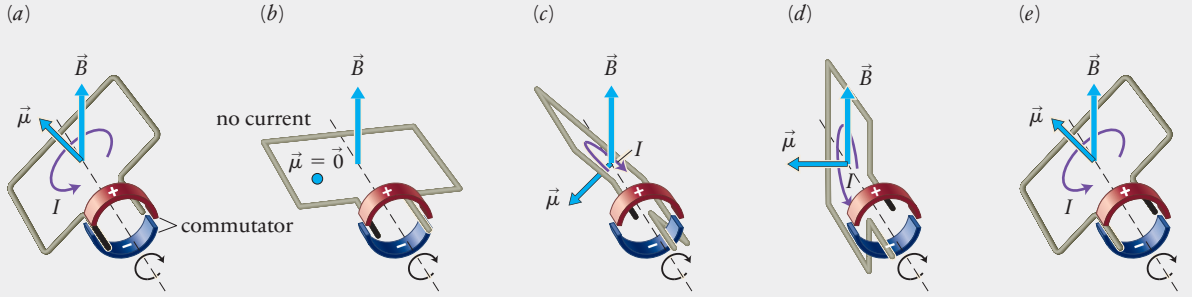
Electric motors

The torque caused by the forces exerted on a current loop in a magnetic field is the basic operating mechanism of an electric motor. A problem with the arrangement shown in Figure 28.12, however, is that once the magnetic dipole moment is aligned with the magnetic field, the torque disappears. Also if the current loop overshoots this equilibrium position, the torque reverses direction (see Checkpoint 28.5*d*). The most common way to overcome this problem is with a *commutator*—an arrangement of two curved plates that reverses the current through the current loop each half turn. The result is that the current loop keeps rotating.

The basic operation is illustrated in **Figure 28.14**. Each half of the commutator is connected to one terminal of a battery and is in contact with one end of the current loop. In the position illustrated in Figure 28.14*a*, the black end is in contact with the negative commutator (that is, the half of the commutator connected to the negative battery terminal). Consequently, the current direction is counterclockwise as seen from above through the current loop, and the magnetic dipole moment $\vec{\mu}$ points up and to the left. The vertical magnetic field therefore causes a torque that turns the loop clockwise to align $\vec{\mu}$ with $\vec{B}$.

Once the loop has rotated to the position shown in Figure 28.14*b*, $\vec{\mu}$ is aligned with $\vec{B}$ but the contact between the ends of the current loop and the commutator is broken, which means there is no current in the loop. The loop overshoots this equilibrium position, but as soon as it does so, the current direction reverses because now the black end of the loop is in contact with the positive commutator (Figure 28.14*c*). The current direction is now clockwise as seen from above, and so $\vec{\mu}$ is reversed. Consequently, the current loop continues to rotate clockwise, as illustrated in Figures 28.14*c–e*. After half a revolution, we reach the initial situation again and the sequence repeats.

Figure 28.14 Operating principle of an electric motor. At instant (*b*), the current direction through the current loop is reversed.



CONCEPTS

28.7 Describe the motion of the current loop in Figure 28.12 if the magnitude of the magnetic field between the poles of the magnet is greater on the left than it is on the right.

28.4 Ampèrian paths

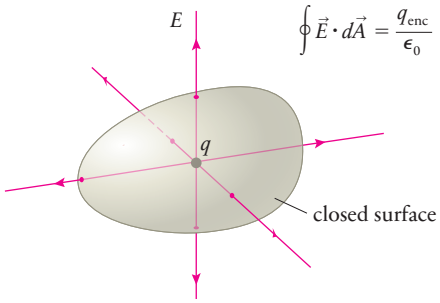
In electrostatics, Gauss's law provides a powerful tool for determining the electric field of a charge distribution: The electric flux through a closed surface is determined by the amount of charge enclosed by that surface (see Section 24.3). The basic reason for Gauss's law is illustrated in **Figure 28.15a**: Electric field lines originate or terminate on charged particles, and the number of "field line piercings" through a closed surface is proportional to the amount of charge enclosed by that surface. As you saw in Checkpoint 27.6, however, Gauss's law for magnetism is not as helpful for determining magnetic fields. The reason is illustrated in Figure 28.15*b*: Magnetic field lines form loops, so if they exit a closed surface, they must reenter it at some other point. Consequently, the magnetic flux through a closed surface is always zero, regardless of whether or not the surface encloses any magnets or current-carrying wires. In mathematical language, the surface integral of the magnetic field over a closed surface is always zero.

Let us next consider line integrals of electric and magnetic fields. As we saw in Section 25.5, the line integral of the electrostatic field around a closed path is always zero. Consider, for example, the line integral of the electrostatic field around closed path 1 in Figure 28.15*c*. Because the electrostatic field generated by the charged particle at the center of path 1 is always perpendicular to the path, $\vec{E}$ is perpendicular to $d\vec{\ell}$. Therefore $\vec{E} \cdot d\vec{\ell} = 0$ and so the line integral is always zero. Any other path, such as closed path 2 in Figure 28.15*c*, can be broken down into small radial and circular segments. As we saw in Section 25.5, going once around a closed path, all the nonzero contributions along the radial segments add up to zero. Physically this means that as you move a charged particle around a closed path through an electrostatic field, the work done by the electrostatic field on the particle is zero.

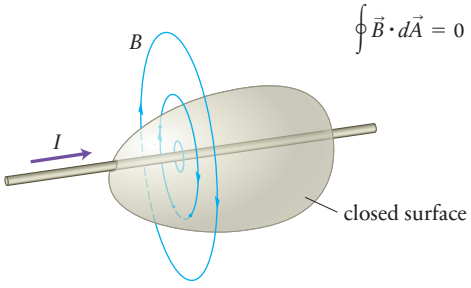
Consider now, however, the line integral around closed path 1 in Figure 28.15*d*, which is concentric with a wire carrying a current out of the page. The magnetic field generated by the current through the wire always points in the same direction as the direction of this path. Therefore $\vec{B} \cdot d\vec{\ell}$ is always positive, making the line integral nonzero and positive. Along closed path 2, the component of the magnetic field tangent to the path is always opposite the direction of the path. Therefore $\vec{B} \cdot d\vec{\ell}$ is always negative and the line integral is negative.

Figure 28.15 Surface and line integrals of electric and magnetic fields.

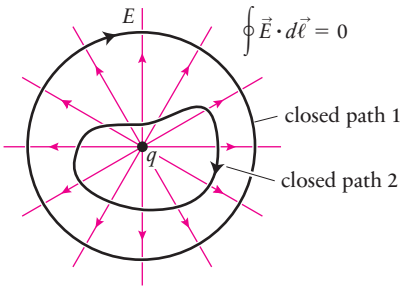
(*a*) Surface integral of electric field (Gauss's law)



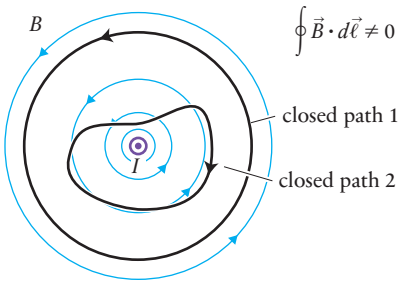
(*b*) Surface integral of magnetic field (Gauss's law for magnetism)



(*c*) Line integral of electric field



(*d*) Line integral of magnetic field



CONCEPTS

Figure 28.16 (*a*) Two closed circular paths concentric with a wire that carries a current directed out of the page. (*b*) A noncircular path encircling the current-carrying wire. The two arcs each represent one-eighth of a circle.

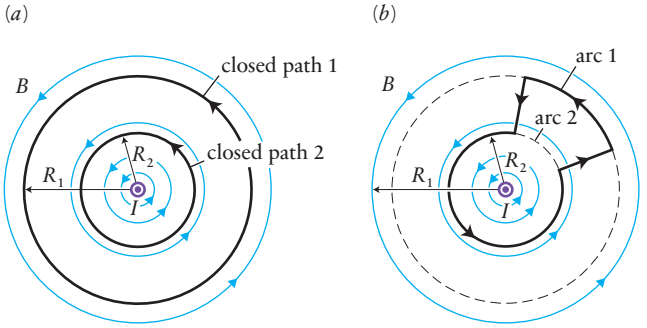
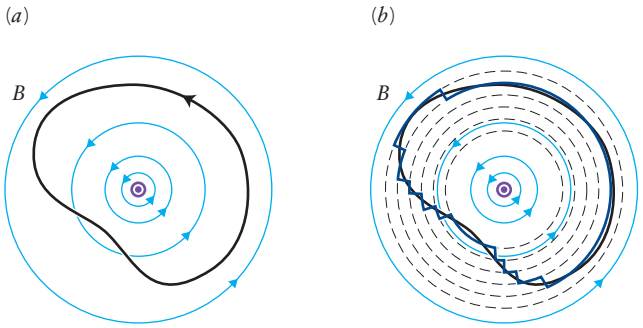


Figure 28.17 (*a*) A noncircular closed path encircling a current-carrying wire. (*b*) We can approximate the path by using small arcs and radial segments.



Let us next compare the line integrals along two closed circular paths of different radius (**Figure 28.16*a***). The arrowheads in the paths indicate the direction along which we carry out the integration. Let the magnitude of the magnetic field a distance R_1 from the wire at the center of the paths be B_1. Because this magnitude is constant along the entire circular path, the line integral along this path is the product of the field magnitude and the length of the path: $B_1(2\pi R_1) = 2\pi B_1 R_1$. Along closed path 2 we obtain $2\pi B_2 R_2$, where B_2 is the magnetic field magnitude a distance R_2 from the wire. As you may suspect from the cylindrical symmetry of the wire, the magnitude of the magnetic field decreases as $1/r$ with distance r from the wire (we'll confirm this dependence in Example 28.6). This means that as we go from R_2 to R_1, the field decreases by a factor R_2/R_1. In other words, $B_1 = B_2(R_2/R_1)$ or $B_1 R_1 = B_2 R_2$. Thus we see that the line integrals along the two paths are equal: $2\pi B_1 R_1 = 2\pi B_2 R_2$. The same argument can be applied to any other closed circular path centered on the wire, from which we conclude that the line integral of the magnetic field of a straight current-carrying wire over any circular path centered on the wire has the same value.

CONCEPTS

28.8 If the magnitude of the current I through a wire is increased, do you expect the line integral of the magnetic field around a closed path around the wire to increase, decrease, or stay the same?

Now consider the noncircular path illustrated in Figure 28.16*b*. Most of the path lies along a circle of radius R_2, with the exception of one-eighth of a revolution, which is along two radial segments and an arc of radius R_1. Because the radial segments are perpendicular to the magnetic field, they do not contribute to the line integral. How do the line integrals along arcs 1 and 2 compare? As we just saw, the line integrals along the two closed circular paths in Figure 28.16*a* are equal, so the line integrals along one-eighth of each closed path in Figure 28.16*a* must also be equal. The same must be true in Figure 28.16*b*, which means the line integrals along arcs 1 and 2 are identical, and so the line integral along the noncircular path in Figure 28.16*b* is equal to that along any circular path centered on the wire.

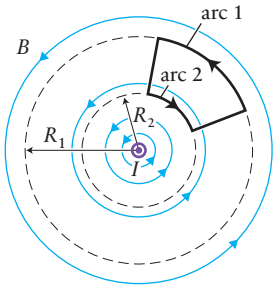
We can make the deviations from a circular path progressively more complicated, but as illustrated in **Figure 28.17**, any path can always be broken down into small segments that are either radial or circular and concentric with the wire. The radial segments never contribute to the line integral because the magnetic field is always perpendicular to them, while the circular segments always add up to a single complete revolution. So, in conclusion:

The value of the line integral of the magnetic field along a closed path encircling a current-carrying wire is independent of the shape of the path.

28.9 What happens to the value of the line integral along the closed path in Figure 28.17*a* when (*a*) the direction of the current through the wire is reversed; (*b*) a second wire carrying an identical current is added parallel to and to the right of the first one (but still inside the path); and (*c*) the direction of the current through the second wire is reversed?

Next let's examine the line integral along a closed path near a current-carrying wire lying outside the path. One such path is shown in **Figure 28.18** as two arcs joined by two radial segments. As before, the radial segments do not contribute to the line integral. The magnitudes of the line integrals along arcs 1 and 2 are equal, but because the direction of arc 2 is opposite the direction of the magnetic field, the line integral along that arc is negative. Consequently, the line integral along the entire path adds up to zero. We can again extend this statement to a path of a different

Figure 28.18 A noncircular closed path not encircling a current-carrying wire.



form, but as long as the path does not encircle the current-carrying wire, the line integral is zero:

> **The line integral of the magnetic field along a closed path that does not encircle any current-carrying wire is zero.**

Putting the above results together, we see that the line integral of the magnetic field along a closed path tells us something about the amount of current encircled by the path:

> **The line integral of the magnetic field along a closed path is proportional to the current encircled by the path.**

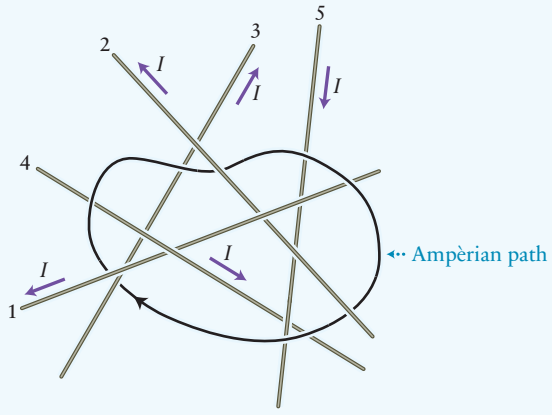
We shall put this statement in a more quantitative form in Section 28.5. As we shall see there, this law, called **Ampère's law,** plays a role analogous to Gauss's law: Given the amount of current encircled by a closed path, called an **Ampèrian path,** we can readily determine the magnetic field due to the current, provided the current distribution exhibits certain simple symmetries. Because the line integral along a closed path depends on the direction of integration, we must always choose a direction along the path when specifying an Ampèrian path. Exercise 28.2 illustrates the importance of the direction of the Ampèrian path.

28.10 Suppose the path in Figure 28.17 were tilted instead of being in a plane perpendicular to the current-carrying wire. Would this tilt change the value of the line integral of the magnetic field around the path?

EXERCISE 28.2 Crossed wires

Consider the Ampèrian path going through the collection of current-carrying wires in **Figure 28.19**. If the magnitude of the current is the same in all the wires, is the line integral of the magnetic field along the Ampèrian path positive, negative, or zero?

Figure 28.19 Exercise 28.2.



SOLUTION For each wire, I must determine whether or not the path encircles the wire and, if it does, whether the component of that wire's magnetic field tangent to the path points in the same direction as the path. I see that wires 1 and 3 go through the path but the other three wires lie either on top of the path or beneath it.

The direction of current through wire 1 is forward out of the plane of the page, so the magnetic field lines around this wire curl counterclockwise—opposite the direction of the Ampèrian path—giving a negative contribution to the line integral. The direction of current through wire 3 is down into the plane of the page, so it yields a positive contribution to the line integral. Because the two currents are equal in magnitude, the contributions to the line integral add up to zero. ✔

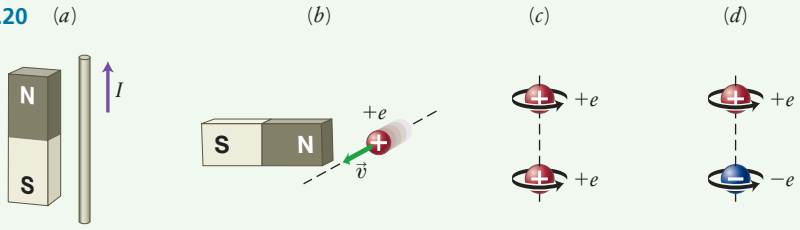
28.11 How do the following changes affect the answer to Exercise 28.2: (*a*) reversing the current through wire 1, (*b*) reversing the current through wire 2, (*c*) reversing the direction of the Ampèrian path?

CONCEPTS

Self-quiz

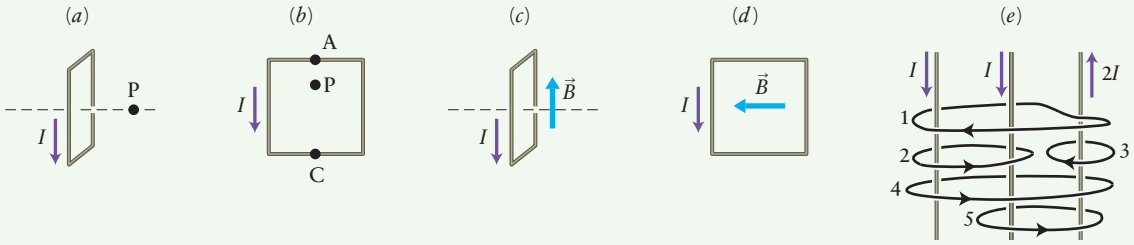
1. Determine the direction of the magnetic force exerted at the center of the wire or on the particles in **Figure 28.20**.

Figure 28.20



2. Determine the direction of the magnetic field at P due to (*a*) the current loop in **Figure 28.21a** and (*b*) segments A and C of the current loop in Figure 28.21*b*.

Figure 28.21

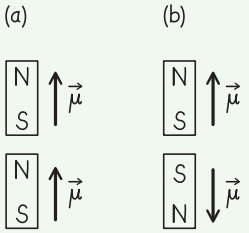


3. Determine in which direction the current loop rotates (*a*) in Figure 28.21*c* and (*b*) in Figure 28.21*d*.
4. (*a*) Determine the currents encircled by the five Ampèrian paths in Figure 28.21*e*. (*b*) Rank the paths according to the magnitudes of the line integral of the magnetic field along each path, greatest first.

Answers

1. (*a*) No magnetic force is exerted by the magnet on the wire because the magnetic field at the location of the wire and the current are antiparallel. (*b*) The magnetic force acting on the particle is upward. To see this, consider the moving positively charged particle to be current in the direction of the velocity of the particle and then use the right-hand force rule, which makes your thumb point upward. (*c*) The magnetic dipole moments of the two spinning particles both point up and so the particles attract each other, just like two bar magnets oriented the way the spinning particles are (**Figure 28.22a**). (*d*) The magnetic dipole moment of the negative particle points down, that of the positive particle points up, and so the two particles repel each other. The comparable bar magnet orientation is shown in Figure 28.22*b*.

Figure 28.22



2. (*a*) The right-hand dipole rule tells you that the magnetic dipole of the loop and the magnetic field produced by the loop point to the right at P. (*b*) Segments A and C both contribute a magnetic field that points out of the page at P, so the magnetic field due to both segments also points out of the page at P.
3. (*a*) $\vec{\mu}$ for the current loop points to the right, so the current loop rotates counterclockwise about an axis perpendicular to the page and through the center of the loop. (*b*) $\vec{\mu}$ for the current loop points out of the page, so the loop rotates about an axis aligned with the vertical sides of the loop. The right side of the loop moves up out of the page, and the left side moves down into the page.
4. (*a*) Path 1 encircles two currents I in the same direction as the magnetic field: $+2I$. Path 2 encircles the same two currents in the opposite direction: $-2I$. Path 3 encircles $2I$ in the direction opposite the magnetic field direction: $-2I$. Path 4 encircles all three currents, which add up to zero. Path 5 encircles I in the direction opposite the magnetic field direction and $2I$ in the same direction as the magnetic field: $+I$. (*b*) Each line integral is proportional to the current encircled, making the ranking line integral 1 = line integral 2 = line integral 3 > line integral 5 > line integral 4.

28.5 Ampère's law

In the preceding section, we saw that the line integral of the magnetic field around a closed path, called an Ampèrian path, is proportional to the current encircled by the path, I_{enc}. This can be expressed mathematically as

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I_{enc} \quad \text{(constant currents)}, \tag{28.1}$$

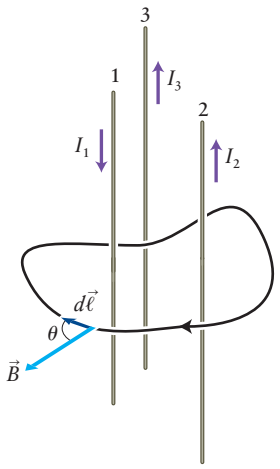
where $d\vec{\ell}$ is an infinitesimal segment of the path and the proportionality constant μ_0 is called the **magnetic constant** (sometimes called *permeability constant*). Using the base unit for current, the ampere (see Section 27.5), the value of the magnetic constant is

$$\mu_0 = 4\pi \times 10^{-7}\ \text{T} \cdot \text{m/A}.$$

Equation 28.1 is called **Ampère's law,** after the French physicist André-Marie Ampère (1775–1836).

Let us illustrate how Eq. 28.1 is used by applying it to the closed path in Figure 28.23. From the figure we see that the path encircles wires 1 and 2 but not 3, which lies behind the path. To calculate the right side of Eq. 28.1, we must assign an algebraic sign to the contributions of currents I_1 and I_2 to I_{enc}. As we saw in Section 28.3, we can do so using the right-hand current rule. Putting the thumb of our right hand in the direction of I_1, we see that the magnetic field of I_1 curls in the same direction as the direction of the integration path (the Ampèrian path) indicated in the diagram. Applying the same rule to I_2 tells us that the magnetic field of this current curls in the opposite direction. Thus I_1 yields a positive contribution and I_2 yields a negative contribution to the line integral. The right side of Eq. 28.1 thus becomes $\mu_0(I_1 - I_2)$.

Figure 28.23 A closed path encircling two of three straight current-carrying wires.



PROCEDURE: Calculating the magnetic field using Ampère's law

For magnetic fields with straight or circular field lines, Ampère's law allows you to calculate the magnitude of the magnetic field without having to carry out any integrations.

1. Sketch the current distribution and the magnetic field by drawing one or more field lines using the right-hand current rule. A two-dimensional drawing should suffice.
2. If the field lines form circles, the Ampèrian path should be a circle. If the field lines are straight, the path should be rectangular.
3. Position the Ampèrian path in your drawing such that the magnetic field is either perpendicular or tangent to the path and constant in magnitude. Choose the direction of the Ampèrian path so that, where it runs parallel to the magnetic field lines, it points in the same direction as the field. If the current distribution divides space into distinct regions, draw an Ampèrian path in each region where you wish to calculate the magnetic field.
4. Use the right-hand current rule to determine the direction of the magnetic field of each current encircled by the path. If this magnetic field and the Ampèrian path have the same direction, the contribution of the current to I_{enc} is positive. If they have opposite directions, the contribution is negative.
5. For each Ampèrian path, calculate the line integral of the magnetic field along the path. Express your result in terms of the unknown magnitude of the magnetic field B along the Ampèrian path.
6. Use Ampère's law (Eq. 28.1) to relate I_{enc} and the line integral of the magnetic field and solve for B. (If your calculation yields a negative value for B, then the magnetic field points in the direction opposite the direction you assumed in step 1.)

You can use the same general approach to determine the current given the magnetic field of a current distribution. Follow the same procedure, but in steps 4–6, express I_{enc} in terms of the unknown current I and solve for I.

QUANTITATIVE TOOLS

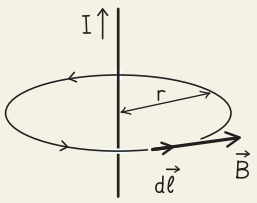
To calculate the line integral on the left side of Eq. 28.1, we first divide the closed path into infinitesimally small segments $d\vec{\ell}$, one of which is shown in Figure 28.23. The segments are directed tangentially along the path in the direction of integration. For each segment $d\vec{\ell}$, we take the scalar product of $d\vec{\ell}$ with the magnetic field $\vec{B}$ at the location of that segment, $\vec{B} \cdot d\vec{\ell}$, and then we add up the scalar products for all segments of the closed path. In the limit that the segment lengths approach zero, this summation becomes the line integral on the left in Eq. 28.1. In the situation illustrated in Figure 28.23, we cannot easily carry out this integration because of the irregular shape of the path. Just like Gauss's law in electrostatics, however, Ampère's law allows us to easily determine the magnetic field for highly symmetrical current configurations. The general procedure is outlined in the Procedure box on the previous page. The next two examples illustrate how this procedure can be applied to simplify calculating the line integral in Ampère's law.

EXAMPLE 28.3 Magnetic field generated by a long straight current-carrying wire

A long straight wire carries a current of magnitude I, and this current creates a magnetic field $\vec{B}$. Derive an expression for the magnitude of the magnetic field a radial distance r from the wire.

❶ **GETTING STARTED** I begin by making a sketch of the wire, arbitrarily orienting it vertically (**Figure 28.24**). I know that the magnetic field is circular, so I draw one circular field line, centering it on the wire and giving it a radius r. Using the right-hand current rule, I determine the direction in which the magnetic field points along the circle and indicate that with an arrowhead in my drawing.

Figure 28.24



❷ **DEVISE PLAN** To begin my derivation I can use Ampère's law (Eq. 28.1). If I let the field line in my sketch be the Ampèrian path, the magnetic field is constant in magnitude and tangential all along the circular path, simplifying the integral on the left in Eq. 28.1. I let the direction of the Ampèrian path be the same as that of the magnetic field. That is to say, the infinitesimal path segment $d\vec{\ell}$ points in the same direction as $\vec{B}$ all along the Ampèrian path.

❸ **EXECUTE PLAN** With my choice of Ampèrian path, $d\vec{\ell}$ and $\vec{B}$ are always pointing in the same direction, and so $\vec{B} \cdot d\vec{\ell} = B\, dl$. Because the magnitude B is the same all around the path, I can write for the left side of Eq. 28.1

$$\oint \vec{B} \cdot d\vec{\ell} = \oint B\, d\ell = B \oint d\ell.$$

The line integral on the right here is the sum of the lengths of all the segments $d\vec{\ell}$ around the Ampèrian path; that is to say, it is equal to the circumference of the circle. Therefore I have

$$\oint \vec{B} \cdot d\vec{\ell} = B \oint d\ell = B(2\pi r). \qquad (1)$$

Because the direction in which the Ampèrian path encircles the current I is the same as the direction of the magnetic field generated by the current, the right side of Eq. 28.1 yields

$$\mu_0 I_{\text{enc}} = +\mu_0 I. \qquad (2)$$

Substituting the right side of Eq. 2 and Eq. 1 into Eq. 28.1, I get $B(2\pi r) = \mu_0 I$ or

$$B = \frac{\mu_0 I}{2\pi r}. \checkmark$$

❹ **EVALUATE RESULT** My result shows that the magnitude of the magnetic field is proportional to I, as I expect (doubling the current should double the magnetic field), and inversely proportional to the radial distance r from the wire. I know from Chapter 24 that the electric field is also inversely proportional to r in cases that exhibit cylindrical symmetry, another indication that my result here makes sense.

QUANTITATIVE TOOLS

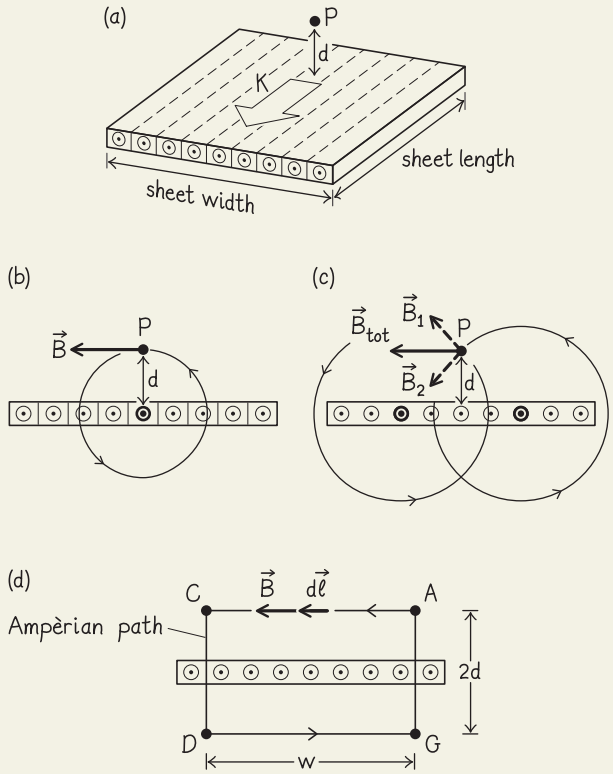
28.12 Suppose the wire in Example 28.3 has a radius R and the current is uniformly distributed throughout the volume of the wire. Follow the procedure of Example 28.3 to calculate the magnitude of the magnetic field inside ($r < R$) and outside ($r > R$) the wire.

EXAMPLE 28.4 Magnetic field generated by a large current-carrying sheet

A large flat metal sheet carries a current. The magnitude of the current per unit of sheet width is K. What is the magnitude of the magnetic field a distance d above the sheet?

❶ **GETTING STARTED** I begin by drawing the sheet and indicating a point P a distance d above it where I am to determine the magnetic field magnitude (**Figure 28.25*a***). I draw a lengthwise arrow to show the current through the sheet.

Figure 28.25



❷ **DEVISE PLAN** I'll solve this problem using Ampère's law (Eq. 28.1). I first need to determine the direction of the magnetic field on either side of the sheet. To this end, I divide the sheet into thin parallel strips, as indicated by the dashed lines in Figure 28.25*a*. I can then treat each strip as a current-carrying wire.

Figure 28.25*b* is a cross-sectional view of the sheet after I have divided it into strips. The perspective here is looking into the sheet in the direction opposite the direction of the current. Using the right-hand current rule, I determine that the strip right underneath point P contributes a magnetic field at P that points parallel to the sheet and to the left. Next I look at the contributions from the two strips labeled 1 and 2 in Figure 28.25*c*, equidistant on either side of P. The two large circles show the magnetic field lines from these two strips that go through P. Because the magnitudes of the contributions from the two strips are equal, the vector sum of their contributions to the magnetic field also points parallel to the sheet and to the left. The same argument can be applied to any other pair of strips that are equidistant on either side of P. Thus, at P the magnetic field due to the entire sheet must be parallel to the sheet and to the left. Similar reasoning shows that the magnetic field below the sheet is also parallel to the sheet and points to the right.

To exploit what I know about the magnetic field, I choose the rectangular path ACDG in Figure 28.25*d* as my Ampèrian path, making the direction of the path the same as that of the magnetic field. The width of this path is w, and its height is $2d$.

❸ **EXECUTE PLAN** I can write the line integral around the Ampèrian path as the sum of four line integrals, one over each side of the path:

$$\oint \vec{B}\cdot d\vec{\ell} = \int_A^C \vec{B}\cdot d\vec{\ell} + \int_C^D \vec{B}\cdot d\vec{\ell} + \int_D^G \vec{B}\cdot d\vec{\ell} + \int_G^A \vec{B}\cdot d\vec{\ell}.$$

Along sides CD and GA the magnetic field is perpendicular to the Ampèrian path and so $\vec{B}\cdot d\vec{\ell}$ is zero. Along each of the two horizontal sides, $d\vec{\ell}$ and $\vec{B}$ point in the same direction, and so $\vec{B}\cdot d\vec{\ell} = B\,d\ell$.

Symmetry requires that the magnitude of the magnetic field a distance d below the sheet be the same as the magnitude a distance d above it (because flipping the sheet upside down does not alter the physical situation). Therefore I can take the magnitude B out of the integral:

$$\oint \vec{B}\cdot d\vec{\ell} = \int_A^C B\,d\ell + \int_D^G B\,d\ell = B\int_A^C d\ell + B\int_D^G d\ell.$$

These two line integrals $\int d\ell$ yield the length of the two horizontal sides, which is the Ampèrian path width w, and so the left side of Eq. 28.1 becomes

$$\oint \vec{B}\cdot d\vec{\ell} = B(w+w) = 2Bw. \qquad (1)$$

To get an expression for the right side of Eq. 28.1, I must first determine the amount of current encircled by the Ampèrian path. Because the magnitude of the current per unit of width through the sheet is K, the magnitude of the current through the Ampèrian path of width w is Kw. The right side of Eq. 28.1 thus becomes

$$\mu_0 I_{\text{enc}} = \mu_0 K w. \qquad (2)$$

Substituting the right sides of Eqs. 1 and 2 into Ampère's law, I get $2Bw = \mu_0 K w$, or

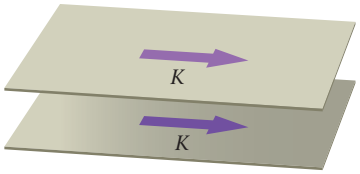
$$B = \tfrac{1}{2}\mu_0 K. \checkmark$$

❹ **EVALUATE RESULT** Because the sheet has planar symmetry (see Section 24.4), I expect that, by analogy with the electric field around a flat charged sheet, the magnetic field on either side of my sheet here is uniform. That is, the field magnitude does not depend on the distance from the sheet, and this is just what my result shows. It also makes sense that my result shows that the magnitude of the magnetic field is proportional to the current per unit width K through the sheet.

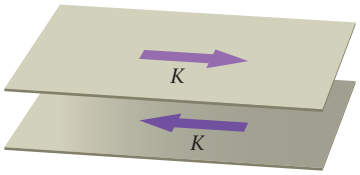
QUANTITATIVE TOOLS

Figure 28.26 Checkpoint 28.13.

(*a*)



(*b*)



28.13 (*a*) What are the direction and magnitude of the magnetic field between the parallel current-carrying sheets of **Figure 28.26a**? What is the direction of $\vec{B}$ outside these sheets? (*b*) Repeat for the sheets of Figure 28.26*b*.

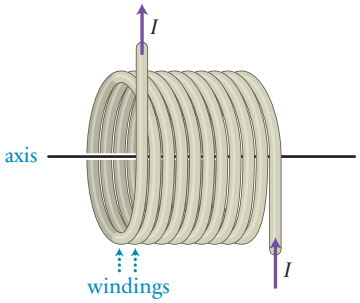
28.6 Solenoids and toroids

A **solenoid** is a long, tightly wound helical coil of wire (**Figure 28.27a**). In general, the diameter of the coil is much smaller than the length of the coil. When a current enters a solenoid at one end and exits at the other end, the solenoid generates a strong magnetic field. If a magnetic core is placed in the solenoid, the solenoid exerts a strong magnetic force on the core, turning electrical energy into motion. Solenoids are therefore often used in electrical valves and actuators. Because solenoids are generally very tightly wound, we can treat the windings of a solenoid as a stack of closely spaced current loops (Figure 28.27*b*).

Figure 28.28a shows that the magnetic field inside a long solenoid must be directed along the axis of the solenoid. Consider, for example, point P inside the solenoid. The figure shows the field lines of two of the loops, one on either side of P and equidistant from that point. The magnetic field contributions of the two loops give rise to a magnetic field that points along the axis of the solenoid. The same argument can be applied to any other pair of loops and to any other point inside the solenoid. Therefore the magnetic field everywhere inside a long solenoid must be directed parallel to the solenoid axis.

Figure 28.27 A solenoid is a tightly wound helical coil of wire.

(*a*) A solenoid



(*b*) A solenoid approximated as a stack of parallel current loops

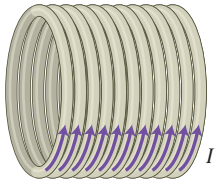
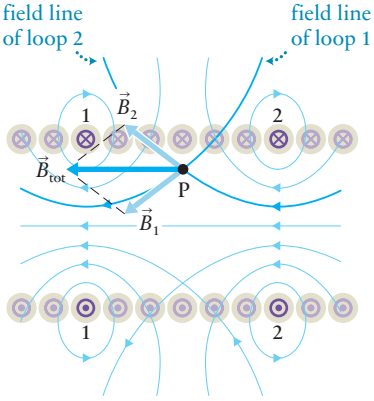
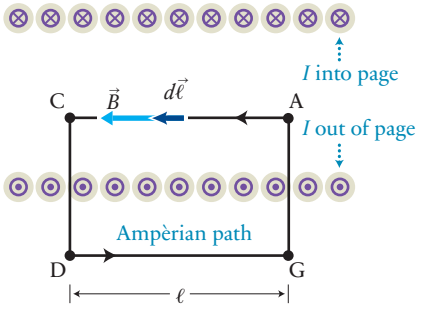


Figure 28.28 The magnetic field of a solenoid that carries a current *I*.

(*a*) Cross section of a solenoid showing the contribution of two loops to the magnetic field at a point inside the solenoid



(*b*) Ampèrian path for calculating the magnetic field inside the solenoid



QUANTITATIVE TOOLS

Because magnetic field lines form loops, all the lines that go through the solenoid must loop back from one end of the solenoid to the other. Because there is much more space available outside the solenoid, the density of the field lines as they loop back is much smaller than the field line density inside the solenoid. The longer the solenoid, the smaller the field line density in the immediate vicinity outside the solenoid. In the limit of an infinitely long solenoid, we can expect the magnetic field outside the solenoid to approach zero.

We can now use Ampère's law to determine the magnitude of the magnetic field when there is a current of magnitude I through the solenoid. Exploiting what we know about the direction of the magnetic field, we choose the rectangular path ACDG in Figure 28.28*b* as the Ampèrian path. Along AC, the magnetic field is parallel to the path, and so $\vec{B} \cdot d\vec{\ell} = B\, d\ell$. Along CD and GA, the magnetic field is perpendicular to the path, so these segments do not contribute: $\vec{B} \cdot d\vec{\ell} = 0$. Because the magnetic field approaches zero outside the solenoid, the segment DG also does not contribute. The line integral on the left side of Ampère's law (Eq. 28.1) thus becomes

$$\oint \vec{B} \cdot d\vec{\ell} = \int_A^C B\, d\ell. \tag{28.2}$$

The cylindrical symmetry of the solenoid requires the magnitude of the magnetic field to be constant along AC, so we can pull B out of the integral:

$$\oint \vec{B} \cdot d\vec{\ell} = \int_A^C B\, d\ell = B \int_A^C d\ell = B\ell, \tag{28.3}$$

where ℓ is the length of side AC.

What is the current encircled by the Ampèrian path? Each winding carries a current of magnitude I, but the path encircles more than one winding. If there are n windings per unit length, then $n\ell$ windings are encircled by the Ampèrian path, and the encircled current is I times the number of windings:

$$I_{\text{enc}} = n\ell I. \tag{28.4}$$

Substituting Eqs. 28.3 and 28.4 into Ampère's law (Eq. 28.1), we obtain

$$B\ell = \mu_0 n\ell I \tag{28.5}$$

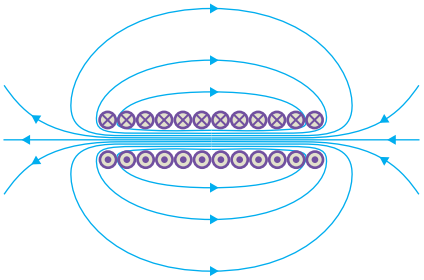
$$B = \mu_0 n I \quad \text{(infinitely long solenoid)}. \tag{28.6}$$

This result shows that the magnetic field inside the solenoid depends on the current through the windings and on the number of windings per unit length. The field within the solenoid is uniform—it does not depend on position inside the solenoid. Although Eq. 28.6 holds for an infinitely long solenoid, the result is pretty accurate even for a solenoid of finite length. For a solenoid that is at least four times as long as it is wide, the magnetic field is very weak outside the solenoid and approximately uniform and equal to the value given in Eq. 28.6 inside. An example of the magnetic field of a finite solenoid is shown in **Figure 28.29**. Note how the magnetic field pattern resembles that of the magnetic field around a bar magnet.

If a solenoid is bent into a circle so that its two ends are connected (**Figure 28.30*a***), we obtain a **toroid**. The magnetic field lines in the interior of a toroid (that is, the donut-shaped cavity enclosed by the coiled wire) close on themselves; thus they do not need to reconnect outside the toroid, as in the case of a solenoid. The entire magnetic field is contained inside

Figure 28.29 Magnetic field line pattern in a solenoid of finite length.

(*a*) Magnetic field of a solenoid, shown in cross section

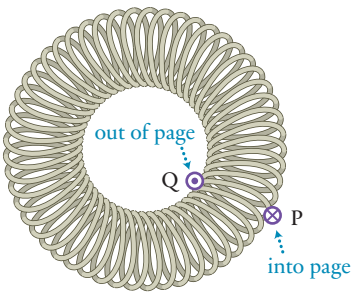


(*b*) The magnetic field of a solenoid

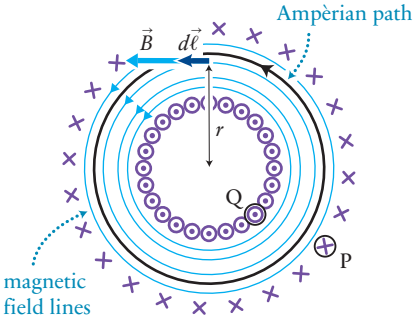


Figure 28.30 A toroid is a solenoid bent into a ring.

(*a*) Toroid



(*b*) Cross section showing magnetic field and a choice of Ampèrian path



QUANTITATIVE TOOLS

the cavity. Symmetry requires the field lines to form circles inside the cavity; the field lines run in the direction of your right thumb when you curl the fingers of your right hand along the direction of the current through the windings. Figure 28.30*b* shows a few representative magnetic field lines in a cross section of the toroid where the current goes into the page on the outside rim of the toroid (as at point P, for example) and out of the page on the inside rim (as at point Q).

To determine the magnitude of the magnetic field, we apply Ampère's law (Eq. 28.1) to a circular path of radius r that coincides with a magnetic field line (Figure 28.30*b*). Because the field is tangential to the integration path, we have $\vec{B} \cdot d\vec{\ell} = B\, d\ell$. Furthermore, symmetry requires the magnitude of the magnetic field to be the same all along the field line, and so we can pull B out of the integration:

$$\oint \vec{B} \cdot d\vec{\ell} = \oint B\, d\ell = B \oint dl = B(2\pi r). \tag{28.7}$$

The Ampèrian path encircles one side of all of the windings, so if there are N windings, the encircled current is

$$I_{\text{enc}} = NI. \tag{28.8}$$

Substituting these last two equations into Ampère's law (Eq. 28.1), we obtain

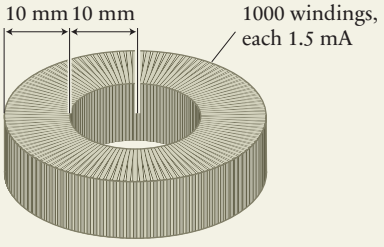
$$B = \mu_0 \frac{NI}{2\pi r} \quad \text{(toroid)}. \tag{28.9}$$

This result tells us that in contrast to a solenoid, the magnitude of the magnetic field in a toroid is not constant—it depends on the distance r to the axis through the center of the toroid.

EXAMPLE 28.5 Square toroid

The toroid in **Figure 28.31** has 1000 windings carrying a current of 1.5 mA. Each winding is a square of side length 10 mm, and the toroid's inner radius is 10 mm. What is the magnitude of the magnetic field at the center of the winding squares?

Figure 28.31 Example 28.5.



❶ **GETTING STARTED** The fact that the windings are square does not change the magnetic field pattern. The magnetic field is still circular as in Figure 28.30*b*, and the arguments given in the derivation of Eq. 28.9 still apply.

❷ **DEVISE PLAN** I can use Eq. 28.9 to determine the magnitude of the magnetic field.

❸ **EXECUTE PLAN** The distance from the center of the toroid to the center of each winding is 10 mm + 5 mm = 15 mm, so

$$B = (4\pi \times 10^{-7}\ \text{T} \cdot \text{m/A}) \frac{(1000)(1.5 \times 10^{-3}\ \text{A})}{2\pi(0.015\ \text{m})}$$

$$= 2.0 \times 10^{-5}\ \text{T}. \checkmark$$

❹ **EVALUATE RESULT** The magnetic field magnitude I obtain is small—comparable to the magnitude of Earth's magnetic field at ground level—but the current through the toroid is very small, so my answer is not unreasonable.

28.14 Use Ampère's law to determine the magnetic field outside a toroid at a distance r from the center of the toroid (*a*) when r is greater than the toroid's outer radius and (*b*) when r is smaller than the toroid's inner radius.

QUANTITATIVE TOOLS

28.7 Magnetic fields due to currents

Ampère's law allows us to determine the magnetic field in only a few symmetrical situations involving current-carrying conductors. For any other situation, such as the one illustrated in **Figure 28.32*a***, Ampère's law is of little help. Suppose, for example, that we are interested in determining the magnetic field at point P due to the current of magnitude I through this conductor. To calculate this field, we develop a procedure that parallels the procedure we developed in Section 23.7 to calculate the electric field of continuous charge distributions.

We begin by treating the wire as a current path and dividing it into small segments, each of length $d\ell$. For each segment we can define a vector $d\vec{\ell}$ that has length $d\ell$ and points in the direction of the current. One such vector segment is shown in Figure 28.32*b*. Let the magnetic field at P due to this segment be $d\vec{B}_s$. If we obtain an expression for the magnetic field due to the segment $d\vec{\ell}$ at an arbitrary position, we can determine the contributions of all the segments that make up the wire to the magnetic field at P and sum these to obtain $\vec{B}$. In the limit of infinitesimally small segments, this summation becomes an integral:

$$\vec{B} = \int_{\substack{\text{current}\\ \text{path}}} d\vec{B}_s, \tag{28.10}$$

where the integration is to be taken along the path followed by the current—that is, a path in the shape of the wire and in the direction of the current.

Before we can carry out the integration, we need to obtain an expression for the magnetic field $d\vec{B}_s$ of a current-carrying segment $d\vec{\ell}$. Because the magnetic field caused by one small segment is too feeble to measure, we cannot determine this field experimentally. From the expression for the magnetic field of a long straight current-carrying wire, however, it is possible to deduce what the field of a segment such as $d\vec{\ell}$ should be.

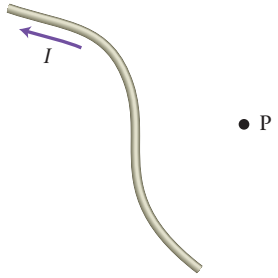
Let us first examine the direction of the magnetic field contribution $d\vec{B}_s$ due to a segment $d\vec{\ell}$ that is located such that $d\vec{\ell}$ is perpendicular to the vector $\vec{r}_{sP}$ pointing from $d\vec{\ell}$ to P ($\theta = 90°$, Figure 28.32*c*). The field lines of this segment are circles centered on the line of motion of the charge carriers causing the current through $d\vec{\ell}$ (see Figures 28.32*a* and *c*). Therefore the magnetic field at P due to $d\vec{\ell}$ is tangent to the circular field line through P. The direction of $d\vec{B}_s$ can be determined by associating the direction of $d\vec{B}_s$ with a vector product. To this end, we denote the unit vector pointing from the segment to P by $\hat{r}_{sP}$, as shown in Figure 28.32*c*. The direction of the vector product $d\vec{\ell} \times \hat{r}_{sP}$ is that of $d\vec{B}_s$: If you curl the fingers of your right hand from $d\vec{\ell}$ to $\hat{r}_{sP}$ in Figure 28.32*c*, your thumb points in the direction of $d\vec{B}_s$.

We expect the magnitude dB_s to be proportional to the current I through the segment and to the length $d\ell$ of the segment. The greater the current or the longer the segment, the stronger the magnetic field. We also expect dB_s to depend on the distance r_{sP} to the segment. The field line picture for magnetic fields suggests that the magnetic field should decrease with distance r_{sP} as $1/r_{sP}^2$, just like the electric field. Finally, for any other segment than the one shown in Figure 28.32*c* we expect the magnetic field to depend on the angle θ between $d\vec{\ell}$ and $\vec{r}_{sP}$: For $\theta = 0$ (which is along the direction of the current) the magnetic field is zero, and for $\theta = 90°$ the magnetic field is maximum. The trigonometric function that fits this behavior is $\sin\theta$. Putting all this information in mathematical form, we obtain

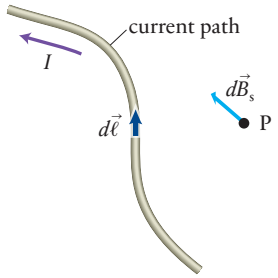
$$dB_s = \frac{\mu_0}{4\pi}\frac{I\,d\ell\sin\theta}{r_{sP}^2} \quad (0 \leq \theta \leq \pi). \tag{28.11}$$

Figure 28.32 The magnetic field at point P due to an arbitrarily shaped current-carrying wire cannot be obtained from Ampère's law. As shown in (*c*), the direction of $d\vec{B}_s$ can be found by taking the vector product of $d\vec{\ell}$, which has length $d\ell$ and direction given by the current through the wire, and $\hat{r}_{sP}$, which points from the segment to point P: $d\vec{\ell} \times \hat{r}_{sP}$.

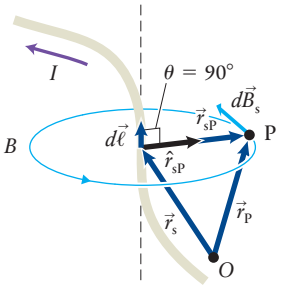
(*a*) Arbitrarily shaped current-carrying wire



(*b*) Small segment $d\vec{\ell}$ contributes magnetic field $d\vec{B}_s$ at point P



(*c*) Determining the direction of $d\vec{B}_s$



QUANTITATIVE TOOLS

The proportionality factor $\mu_0/4\pi$ is obtained by deriving this expression from Ampère's law, but the mathematics is beyond the scope of this book. Instead, we use the reverse approach and show in Example 28.6 that Eq. 28.11 yields the correct result for a long straight current-carrying wire.

Incorporating what we know about the direction of $d\vec{B}_s$, we can also write Eq. 28.11 in vector form:

$$d\vec{B}_s = \frac{\mu_0}{4\pi}\frac{I\,d\vec{\ell}\times\hat{r}_{sP}}{r_{sP}^2} \quad \text{(constant current)}, \tag{28.12}$$

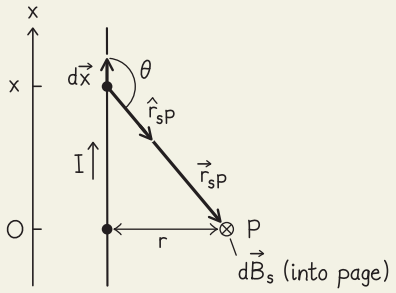
where $\hat{r}_{sP}$ is the unit vector pointing along $\vec{r}_{sP}$ from s to P. Equation 28.12 is known as the **Biot-Savart law**. By substituting Eq. 28.12 into Eq. 28.10, we have a prescription for calculating the magnetic field produced by any constant current.

EXAMPLE 28.6 Another look at the magnetic field generated by a long straight current-carrying wire

A long straight wire carries a current of magnitude I. Use the Biot-Savart law to derive an expression for the magnetic field $\vec{B}$ produced at point P a radial distance r from the wire.

❶ **GETTING STARTED** I begin by making a sketch of the wire (**Figure 28.33**). I arbitrarily orient the wire vertically and then choose the positive x axis along the direction of the wire. Because the magnetic field produced by the current has cylindrical symmetry, I can set the origin anywhere along the axis without loss of generality. For simplicity, I let the origin be at the height of point P. I assume the wire is of infinite length.

Figure 28.33



❷ **DEVISE PLAN** To use the Biot-Savart law (Eq. 28.12), I need to divide the wire into segments, determine the magnetic fields due to all segments, and then take the sum of the fields. In the limit of infinitesimally small segments, this sum becomes the integral given by Eq. 28.10 with the wire serving as the current path along which the integration is carried out.

I indicate one such segment in my sketch, calling it $d\vec{x}$. The magnetic field $d\vec{B}_s$ at P generated by this segment is given by Eq. 28.12. The unit vector $\hat{r}_{sP}$ in this equation points from $d\vec{x}$ to P, and the direction of $d\vec{B}_s$ is given by the right-hand current rule: I point my right thumb along the direction of the current and curl my fingers around the wire to determine the direction of $d\vec{B}_s$ (into the page at P in my sketch). Alternatively, I can use the vector product $d\vec{x}\times\hat{r}_{sP}$ to determine the direction of $d\vec{B}_s$: I line up the fingers of my right hand along $d\vec{x}$ in Figure 28.33 and curl them toward $\hat{r}_{sP}$. When I do this, my thumb points in the direction of the magnetic field. Both methods yield the same result: $d\vec{B}_s$ points into the page.

Note that all the segments $d\vec{x}$ along the wire produce a magnetic field in the same direction. This means I can take the algebraic sum of the *magnitudes* of $d\vec{B}_s$ to determine the magnitude of the magnetic field at P. Then I can use Eq. 28.11 to express dB_s in terms of dx and integrate the resulting expression from $x = -\infty$ to $x = +\infty$ to determine the magnitude of the magnetic field at P.

❸ **EXECUTE PLAN** Because r_{sP} and θ in Eq. 28.11 both depend on x, I need to express them in terms of x before I can carry out the integration. By the Pythagorean theorem,

$$r_{sP}^2 = x^2 + r^2,$$

and remembering that $\sin\theta = \sin(180^\circ - \theta)$, I write

$$\sin\theta = \frac{r}{r_{sP}} = \frac{r}{\sqrt{x^2+r^2}}.$$

Substituting these last two results into Eq. 28.11 and using dx in place of $d\ell$, I get

$$dB_s = \frac{\mu_0 I}{4\pi}\frac{r}{\sqrt{x^2+r^2}}\frac{1}{x^2+r^2}\,dx = \frac{\mu_0 I r}{4\pi}\frac{dx}{[x^2+r^2]^{3/2}},$$

and integrating this result over the length of the wire gives me

$$B = \frac{\mu_0 I r}{4\pi}\int_{-\infty}^{+\infty}\frac{dx}{[x^2+r^2]^{3/2}} = \frac{\mu_0 I r}{4\pi}\left[\frac{1}{r^2}\frac{x}{[x^2+r^2]^{1/2}}\right]_{x=-\infty}^{x=+\infty}$$

$$B = \frac{\mu_0 I}{2\pi r}. \checkmark$$

❹ **EVALUATE RESULT** This is identical to the result I obtained using Ampère's law in Example 28.3, a strong indication that my result here is correct.

QUANTITATIVE TOOLS

28.15 Imagine a long straight wire of semi-infinite length, extending from $x = 0$ to $x = +\infty$, carrying a current of constant magnitude I. What is the magnitude of the magnetic field at a point P located a perpendicular distance d from the end of the wire that is at $x = 0$?

The expression for the magnitude of the magnetic field generated by a long straight wire obtained in Example 28.6 can lead us to an expression for the forces exerted by two current-carrying wires on each other. Consider the situation illustrated in **Figure 28.34**: Two parallel wires, of length ℓ and separated by a distance d, carry currents of magnitudes I_1 and I_2. To determine the magnetic force exerted by wire 1 on wire 2, we must first determine the magnitude and direction of the magnetic field generated by wire 1 at the location of wire 2 and then substitute this information into Eq. 27.8, which gives the magnetic force exerted on a straight current-carrying wire of length ℓ in a magnetic field $\vec{B}$:

$$\vec{F}^B = I\vec{\ell} \times \vec{B}, \tag{28.13}$$

where $\vec{\ell}$ is a vector whose magnitude is given by the length ℓ of the wire and whose direction is given by that of the current through the wire. Using the right-hand current rule, we see that the magnetic field $\vec{B}_1$ generated by wire 1 at the location of wire 2 points into the page (Figure 28.34). The result I obtained in Example 28.6 gives the magnitude of this field:

$$B_1 = \frac{\mu_0 I_1}{2\pi d}. \tag{28.14}$$

Because the magnetic field is perpendicular to $\vec{\ell}$, Eq. 28.13 yields for the magnitude of the magnetic force acting on wire 2

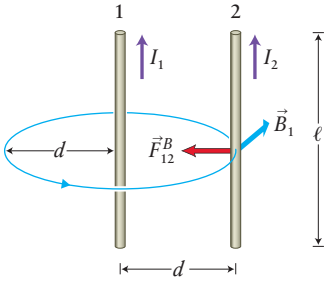
$$F_{12}^B = I_2 \ell B_1 \tag{28.15}$$

or, substituting Eq. 28.14,

$$F_{12}^B = \frac{\mu_0 \ell I_1 I_2}{2\pi d} \quad \text{(parallel straight wires carrying constant currents).} \tag{28.16}$$

The direction of this force follows from the vector product in Eq. 28.13: Using the right-hand force rule (place the fingers of your right hand along $\vec{\ell}$ in Figure 28.34, which points in the direction of I_2, and bend them toward $\vec{B}_1$), you can verify that the force points toward wire 1.

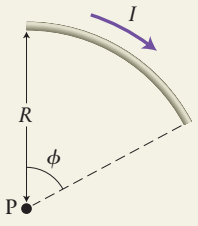
Figure 28.34 Calculating the magnetic force exerted by one current-carrying wire on another.



EXAMPLE 28.7 The magnetic field generated by a circular arc of current-carrying wire

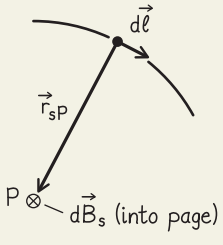
A wire bent into a circular arc of radius R subtending an angle ϕ carries a current of magnitude I (**Figure 28.35**). Use the Biot-Savart law to derive an expression for the magnitude of the magnetic field $\vec{B}$ produced at point P, located at the center of the arc.

Figure 28.35 Example 28.7.



❶ **GETTING STARTED** I begin by evaluating what a small segment of length $d\ell$ along the arc contributes to the magnetic field. I therefore make a sketch showing one segment $d\vec{\ell}$ and the vector $\vec{r}_{sP}$ pointing from the segment to P (**Figure 28.36**). Using the right-hand vector product rule, I see that the direction of $d\vec{\ell} \times \vec{r}_{sP}$, and therefore the magnetic field $d\vec{B}_s$, are into the page at P.

Figure 28.36



(Continued)

QUANTITATIVE TOOLS

2 DEVISE PLAN Because all segments contribute a magnetic field in the same direction, I can integrate Eq. 28.11 over the arc to obtain the magnitude of the magnetic field at P.

3 EXECUTE PLAN Because $d\vec{\ell}$ and $\vec{r}_{sP}$ are always perpendicular to each other, I can write for Eq. 28.11

$$dB_s = \frac{\mu_0}{4\pi}\frac{I\,d\ell\,\sin 90°}{R^2} = \frac{\mu_0}{4\pi}\frac{I\,d\ell}{R^2}, \qquad (1)$$

where I have substituted the radius R for the magnitude of $\vec{r}_{sP}$. To change the integration variable from $d\ell$ to the angular variable $d\phi$, I substitute $d\ell = R\,d\phi$ into Eq. 1:

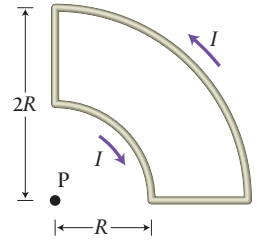
$$dB_s = \frac{\mu_0 I}{4\pi R}\,d\phi$$

$$B = \int_{arc} dB_s = \frac{\mu_0 I}{4\pi R}\int_0^{\phi} d\phi$$

$$= \frac{\mu_0 I\phi}{4\pi R}. \checkmark$$

4 EVALUATE RESULT My expression for B shows that at P the magnetic field magnitude is proportional to the current through the arc, as I expect. It is also proportional to the angle ϕ of the arc (when measured in radians), also what I expect given that two such arcs should yield twice the magnetic field. Finally, B is inversely proportional to the arc radius R. That dependence on the radius makes sense because increasing the radius increases the distance between P and the arc, thus diminishing the magnetic field.

Figure 28.37 A current loop carrying a current of constant magnitude I (Checkpoint 28.16).



28.16 What is the magnitude of the magnetic field (*a*) at the center of a circular current loop of radius R and (*b*) at point P near the current loop in Figure 28.37? Both arcs carry a current of constant magnitude I.

28.8 Magnetic field of a moving charged particle

Let us now use the Biot-Savart law to obtain an expression for the magnetic field caused by charged particles moving at constant velocity.* Consider first a straight wire carrying a current of magnitude I and aligned with the x axis, as in Figure 28.38*a*. The magnetic field generated at point P by a small segment $d\vec{x}$ is given by Eq. 28.12:

$$d\vec{B}_s = \frac{\mu_0}{4\pi}\frac{I\,d\vec{x}\times\hat{r}_{sP}}{r_{sP}^2}, \qquad (28.17)$$

where $\hat{r}_{sP}$ is a unit vector pointing from the segment $d\vec{x}$ to the point at which we wish to determine the magnetic field, and r_{sP} is the distance between the segment and the point P. Suppose the segment contains an amount of charge dq. Let the charge carriers responsible for the current take a time interval dt to have displacement $d\vec{x}$ (Figure 28.38*b*). According to the definition of current (Eq. 27.2), we have

$$I \equiv \frac{dq}{dt}, \qquad (28.18)$$

and so

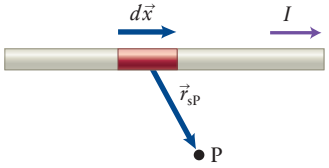
$$Id\vec{x} = \frac{dq}{dt}\,d\vec{x} = dq\,\frac{d\vec{x}}{dt} = dq\,\vec{v}, \qquad (28.19)$$

where $\vec{v}$ is the velocity at which the charge carriers move down the wire. In the limiting case where the segment $d\vec{x}$ contains just a single charge carrier carrying a charge q (Figure 28.38*c*), dq becomes q and

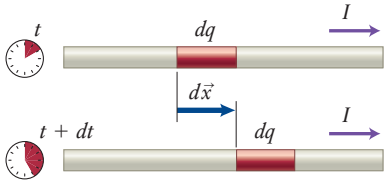
$$Id\vec{x} = q\vec{v}. \qquad (28.20)$$

Figure 28.38 We use the Biot-Savart law to obtain an expression for the magnetic field caused by charged particles moving at constant velocity.

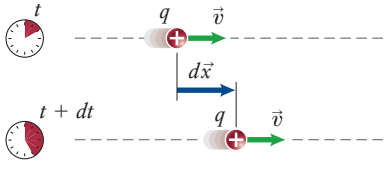
(*a*) Small segment of current-carrying wire causes magnetic field at point P



(*b*) Displacement of charge dq in time interval dt



(*c*) Displacement of charged particle in time interval dt



*In the derivation that follows, we assume $v \ll c_0$ and ignore any relativistic effects as described in Sections 27.4 and 27.8.

QUANTITATIVE TOOLS

Substituting this result into Eq. 28.17, we obtain an expression for the magnetic field of a single moving charged particle:

$$\vec{B} = \frac{\mu_0}{4\pi}\frac{q\vec{v}\times\hat{r}_{\mathrm{pP}}}{r_{\mathrm{pP}}^2} \quad \text{(single particle)}, \tag{28.21}$$

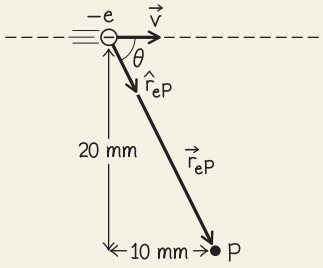
where r_{pP} is the distance between the particle and P, and $\hat{r}_{\mathrm{pP}}$ is the unit vector pointing from the particle to P.

EXAMPLE 28.8 Magnetic field generated by a moving electron

An electron carrying a charge $-e = -1.60\times10^{-19}$ C moves in a straight line at a speed $v = 3.0\times10^{7}$ m/s. What are the magnitude and direction of the magnetic field caused by the electron at a point P 10 mm ahead of the electron and 20 mm away from its line of motion?

❶ **GETTING STARTED** I begin by drawing the moving electron and the point P at which I am to determine the magnetic field (**Figure 28.39**).

Figure 28.39



❷ **DEVISE PLAN** The magnetic field created by a moving charged particle is given by Eq. 28.21.

❸ **EXECUTE PLAN** The unit vector $\hat{r}_{\mathrm{eP}}$ points from the electron to P, and so $\vec{v}\times\hat{r}_{\mathrm{eP}}$ points into the page. Because the charge of the electron is negative, the magnetic field points in the opposite direction, out of the page. ✔

The magnitude of the vector $\vec{r}_{\mathrm{eP}}$ in Figure 28.39 is $r_{\mathrm{eP}} = \sqrt{(10\text{ mm})^2+(20\text{ mm})^2} = 22.36$ mm, and $\sin\theta = (20\text{ mm})/(22.36\text{ mm}) = 0.8944$. Substituting these values into Eq. 28.11 thus yields for the magnitude of the magnetic field

$$B = \frac{(4\pi\times10^{-7}\,\mathrm{T\cdot m/A})}{4\pi}\frac{(1.60\times10^{-19}\,\mathrm{C})(3.0\times10^{7}\,\mathrm{m/s})(0.8944)}{(22.36\times10^{-3}\,\mathrm{m})^2}$$

$$= 8.6\times10^{-16}\,\mathrm{T}. \checkmark$$

❹ **EVALUATE RESULT** The magnetic field magnitude I obtained is much too small to be detected, but that's what I expect for a single electron. I can verify the direction of the magnetic field by applying the right-hand current rule to the current caused by the electron. Because the electron carries a negative charge, its motion to the right in Figure 28.39 causes a current to the left. Pointing my right-hand thumb to the left in the figure, I observe that the fingers curl out of the page at P, in agreement with what I determined earlier for the direction of $\vec{B}$.

We can now combine the expression for the magnetic field caused by a moving charged particle (Eq. 28.21) with that for the magnetic force exerted on another moving charged particle (Eq. 27.19; $\vec{F}_{\mathrm{p}}^{B} = q\vec{v}\times\vec{B}$) to determine the magnetic interaction between two moving charged particles. Consider the situation illustrated in **Figure 28.40**. The magnetic field caused by particle 1 at the location of particle 2 is given by Eq. 28.21:

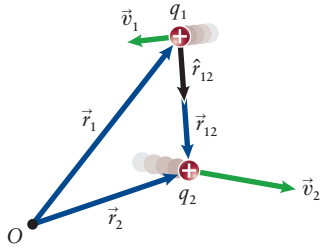
$$\vec{B}_1(\vec{r}_2) = \frac{\mu_0}{4\pi}\frac{q_1\vec{v}_1\times\hat{r}_{12}}{r_{12}^2}, \tag{28.22}$$

where r_{12} is the magnitude of the vector $\vec{r}_{12}$ pointing from particle 1 to particle 2 and $\hat{r}_{12}$ is a unit vector pointing along $\vec{r}_{12}$. Substituting this expression into Eq. 27.19, we obtain for the magnetic force exerted by particle 1 on particle 2

$$\vec{F}_{12}^{B} = q_2\vec{v}_2\times\vec{B}_1(\vec{r}_2) = \frac{\mu_0}{4\pi}\frac{q_1q_2}{r_{12}^2}\vec{v}_2\times(\vec{v}_1\times\hat{r}_{12}). \tag{28.23}$$

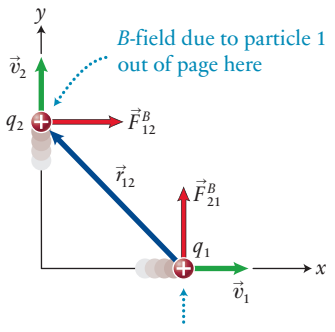
Notice the appearance of the double vector product: We must first take the vector product of $\vec{v}_1$ and $\hat{r}_{12}$ and then take the vector product of $\vec{v}_2$ and the result of the first vector product.

Figure 28.40 Magnetic interaction of two moving charged particles.



QUANTITATIVE TOOLS

Figure 28.41 The forces that moving charged particles exert on each other do not necessarily point in opposite directions.



Equation 28.23 also shows that we need a moving charged particle in order to generate a magnetic field ($\vec{v}_1 \neq 0$) and another moving charged particle to "feel" that magnetic field ($\vec{v}_2 \neq 0$), in agreement with our discussion in the first part of this chapter.

In contrast to the electric force, the magnetic force does not satisfy Eq. 8.15, $\vec{F}_{12} = -\vec{F}_{21}$. To see this, consider the two positively charged moving particles shown in **Figure 28.41**. Particle 1, carrying a charge q_1, travels in the positive x direction, while particle 2, carrying charge q_2, travels in the positive y direction. The force exerted by 1 on 2 is given by Eq. 28.23. Applying the right-hand vector product rule, we determine that the vector product $\vec{v}_1 \times \hat{r}_{12}$ on the right side of Eq. 28.23 points out of the page. Applying the right-hand vector product rule again to the vector product of $\vec{v}_2$ and $\vec{v}_1 \times \hat{r}_{12}$, we see that $\vec{F}_{12}^{B}$ points in the positive x direction.* The force exerted by 2 on 1 is obtained by switching the subscripts 1 and 2 in Eq. 28.23:

$$\vec{F}_{21}^{B} = \frac{\mu_0}{4\pi}\frac{q_1 q_2}{r_{12}^2}\vec{v}_1 \times (\vec{v}_2 \times \hat{r}_{21}), \tag{28.24}$$

where we have used that $q_1 q_2 = q_2 q_1$ and $r_{12}^2 = r_{21}^2$. Applying the right-hand vector product rule twice, we find that $\vec{F}_{21}^{B}$ points in the positive y direction, and so the magnetic forces that the two particles exert on each other, while equal in magnitude, do not point in opposite directions, as we would expect from Eq. 8.15.

We derived Eq. 8.15 from the fact that the momentum of an isolated system of two particles is constant, which in turn follows from conservation of momentum, one of the most fundamental laws of physics. The electric and magnetic fields of isolated moving charged particles, however, are not constant, and as we shall see in Chapter 30, we can associate a flow of both momentum and energy with changing electric and magnetic fields. Therefore we need to account not only for the momenta of the two charged particles in Figure 28.41, but also for the momentum carried by their fields—which goes beyond the scope of this book. Even though Eq. 8.15 is not satisfied by the magnetic force, the momentum of the system comprising the two particles and their fields is still constant.

Substituting Coulomb's law and Eq. 28.23 into Eq. 27.20, $\vec{F}_{\mathrm{p}}^{EB} = q(\vec{E} + \vec{v} \times \vec{B})$, we obtain an expression for the electromagnetic force that two moving charged particles exert on each other:

$$\vec{F}_{12}^{EB} = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r_{12}^2}[\hat{r}_{12} + \mu_0\epsilon_0\vec{v}_2 \times (\vec{v}_1 \times \hat{r}_{12})]. \tag{28.25}$$

By comparing the result obtained in Example 28.6 with Eq. 27.38 and substituting $k = 1/(4\pi\epsilon_0)$ (Eq. 24.7), we see that $\mu_0 = 1/(\epsilon_0 c_0^2)$ or

*Here is another way you can determine the direction of the magnetic force exerted by particle 1 on particle 2. The motion of particle 1 corresponds to a current directed in the positive x direction. Using the right-hand current rule, it follows that the magnetic field of 1 points out of the xy plane at any location above the x axis. The magnetic field due to 1 at the location of 2 thus points out of the plane. The motion of particle 2 corresponds to a current directed in the positive y direction. To determine the direction of the magnetic force exerted on this "current," we can use the right-hand force rule. When you curl the fingers of your right hand from the positive y axis to the direction of $\vec{B}_1$ at particle 2 (out of the page), your thumb points in the positive x direction, as we found before.

QUANTITATIVE TOOLS

$\mu_0\epsilon_0 = 1/c_0^2$, where c_0 is the speed of light. Using this information, we can write Eq. 28.25 as

$$\vec{F}_{12}^{EB} = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r_{12}^2}\left[\hat{r}_{12} + \frac{\vec{v}_2 \times (\vec{v}_1 \times \hat{r}_{12})}{c_0^2}\right] \quad (v \ll c_0). \qquad (28.26)$$

This equation is the most general expression for the electromagnetic interaction between moving charged particles. Because in most applications the speeds of charged particles are significantly smaller than the speed of light, the second term in the square brackets in Eq. 28.26 is much smaller than the first term, which represents the electric contribution to the electromagnetic force from Coulomb's law.



28.17 Consider two protons 1 and 2, each carrying a charge $+e = 1.6 \times 10^{-19}$ C, separated by 1.0 mm moving at 3×10^5 m/s parallel to each other and perpendicular to their separation. (*a*) What is the direction of the magnetic force that each proton exerts on the other? (*b*) Determine the ratio of the magnitudes of the magnetic and electric forces that the two exert on each other.

Chapter Summary

Current loops (Sections 28.1–28.3)

Concepts A magnetic field is produced by magnets, by current-carrying wires, and by moving charged particles.

The magnetic field pattern produced by a wire carrying a current and bent into a circular **current loop** is similar to the field pattern of a magnetic dipole.

A spinning charged particle has a magnetic field identical to that of an infinitesimally small magnetic dipole.

The **magnetic dipole moment** $\vec{\mu}$ is a vector that specifies the orientation and magnitude of a magnetic dipole. The direction of $\vec{\mu}$ is the same as the direction of the magnetic field at the dipole center. For a bar magnet, the direction of $\vec{\mu}$ is from south pole to north pole. For a current loop, the direction of $\vec{\mu}$ is the direction in which your thumb points when you curl the fingers of your right hand in the direction of the current.

A current loop placed in a magnetic field tends to rotate such that the direction of $\vec{\mu}$ aligns with the direction of the magnetic field.

Ampère's law (Sections 28.4, 28.5)

Concepts **Ampère's law** states that the line integral of the magnetic field along a closed path, called an **Ampèrian path,** is proportional to the current encircled by the path. See the Procedure box "Calculating the magnetic field using Ampère's law" on page 987 for how to apply Ampère's law.

Quantitative Tools If a constant current I_{enc} is encircled by an **Ampèrian path,** then by **Ampère's law,** the line integral of the magnetic field around the path is

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I_{\text{enc}}, \tag{28.1}$$

where $d\vec{\ell}$ is an infinitesimal segment of length of the Ampèrian path and μ_0, called the **permeability constant,** has the value

$$\mu_0 = 4\pi \times 10^{-7}\ \text{T} \cdot \text{m/A}.$$

If a long, straight wire carries a current I, the magnitude of the magnetic field it produces at a distance d from the wire is

$$B = \frac{\mu_0 I}{2\pi d}.$$

The magnitude of the magnetic field produced by a large sheet carrying a uniformly distributed current is

$$B = \tfrac{1}{2}\mu_0 K,$$

where K is the current magnitude per unit width of the sheet.

Solenoids and toroids (Section 28.6)

Concepts A **solenoid** is a long, tightly wound helical coil of wire. The diameter of the coil is much smaller than its length. Outside an infinitely long solenoid, the magnetic field approaches zero; inside the solenoid, the magnetic field is directed parallel to the solenoid length.

A **toroid** is a solenoid bent into a circle. The entire magnetic field is contained within the windings of the toroid, and the magnetic field lines form circles around the center of the toroid.

Quantitative Tools If an infinitely long **solenoid** that has n windings per unit length carries a current I, the magnetic field the solenoid produces in its interior has magnitude

$$B = \mu_0 n I. \tag{28.6}$$

If a **toroid** that has N windings carries a current I, the magnetic field the toroid produces inside the windings at a distance r from the toroid center has magnitude

$$B = \mu_0 \frac{NI}{2\pi r}. \tag{28.9}$$

Biot-Savart law (Section 28.7)

Quantitative Tools **Biot-Savart law:** If a small segment of wire $d\vec{\ell}$ carries a constant current I, the magnetic field produced by the segment at a point P a distance r_{sP} from the segment is

$$d\vec{B}_s = \frac{\mu_0}{4\pi} \frac{I\, d\vec{\ell} \times \hat{r}_{sP}}{r_{sP}^2}, \tag{28.12}$$

where $d\vec{\ell}$ points in the current direction and $\hat{r}_{sP}$ points from the segment to P. The magnetic field is

$$\vec{B} = \int_{\text{current path}} d\vec{B}_s. \tag{28.10}$$

A circular arc of radius R spanning an angle ϕ and carrying a current I produces a magnetic field at the center of the circle of magnitude

$$B = \frac{\mu_0 I \phi}{4\pi R}.$$

Magnetic field of a moving charged particle (Section 28.8)

Quantitative Tools A particle that has charge q and velocity $\vec{v}$ produces at a point P a magnetic field

$$\vec{B} = \frac{\mu_0}{4\pi} \frac{q\vec{v} \times \hat{r}_{pP}}{r_{pP}^2} \tag{28.21}$$

where $\hat{r}_{pP}$ points from the particle to P and r_{pP} is the distance from the particle to P. If charged particles 1 and 2 carrying charges q_1 and q_2 are a distance r_{12} apart and have velocities $\vec{v}_1$ and $\vec{v}_2$, the electromagnetic force exerted by particle 1 on particle 2 is

$$\vec{F}_{12}^{EB} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}^2} \left[\hat{r}_{12} + \frac{\vec{v}_2 \times (\vec{v}_1 \times \hat{r}_{12})}{c_0^2} \right], \tag{28.26}$$

where $\hat{r}_{12}$ points from particle 1 to particle 2 and

$$c_0 = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$$

is the speed of light. Equations 28.21 and 28.26 are valid only if $v \ll c_0$.

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = easy, •• = intermediate, ••• = hard; CR = context-rich problem.

28.1 Source of the magnetic field

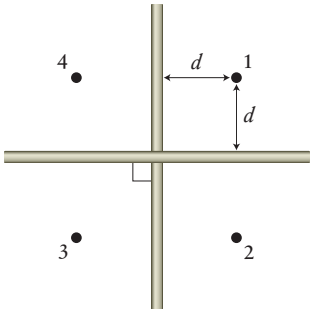
1. A negatively charged particle sits midway between the two magnets in Figure P28.1, at rest relative to the magnets. If the magnet on the left is twice as strong as the magnet on the right, what is the direction of the magnetic force exerted on the particle? •

Figure P28.1



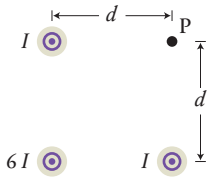
2. There are two parallel wires, each carrying a current I in such a way that in one wire the current flows along positive x-direction and in the other in the negative x-direction. What is the direction of magnetic field at any point at the midway between the two wires and the plane defined by them? •
3. Consider a conducting wire having a shape of a hexagon carries a current that flows in an anti-clockwise manner when viewed from the above. What is the direction of the magnetic field at the center of the hexagon due to each side of the loop? •
4. A negatively charged particle is at rest in a region where a uniform magnetic field points in the positive x direction and a uniform electric field points in the positive y direction. What is the direction of the vector sum of the forces exerted on the particle? ••
5. Suppose two electrons move on parallel, closely spaced paths in the $+z$ direction, each with velocity $\vec{v}$ in the Earth reference frame. Discuss all forces exerted on the electrons. ••
6. Two protons are fired toward each other on closely spaced paths, one moving in the $+z$ direction and one in the $-z$ direction. As they pass close to each other, is the magnetic force between them attractive, repulsive, or neither? ••
7. The two insulated, current-carrying wires in Figure P28.7 cross at right angles, and each carries a current I. The locations labeled 1–4 are all in the plane defined by the wires, with each location a perpendicular distance d from both wires. At how many of these four locations is the magnetic field (*a*) directed out of the page, (*b*) directed into the page, and (*c*) equal to zero? ••

Figure P28.7



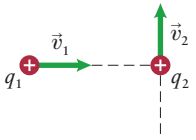
8. Two of the three wires in Figure P28.8 carry current I, and the third wire carries current $6I$. The location labeled P is a distance d from each of the wires carrying current I. What is the direction of the magnetic field at P? ••

Figure P28.8



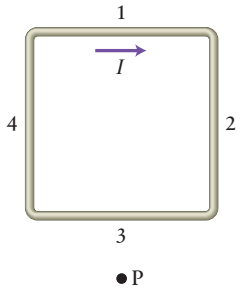
9. Two positively charged particles 1 and 2 are moving in the same plane, with the velocity of particle 1 perpendicular to the velocity of particle 2. At the instant shown in Figure P28.9, particle 2 is on the line that defines the velocity of particle 1. (Take their speeds to be much less than c_0.) For this instant, describe and sketch the direction of the magnetic force that each particle exerts on the other. •••

Figure P28.9



10. In Figure P28.10, (*a*) specify the direction of the magnetic field at the location labeled P due to each of the four sides 1–4 of the current loop. (*b*) Which side produces the strongest magnetic field at P? (*c*) What additional information is needed to determine the direction of the magnetic field at P due to the loop? •••

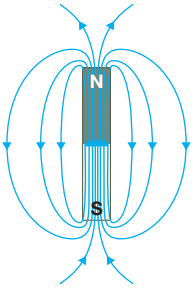
Figure P28.10



28.2 Current loops and spin magnetism

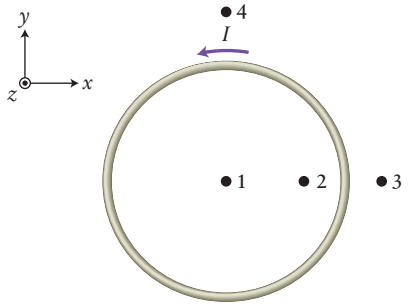
11. A hexagonal wire loop with side 5 cm carries a current 2 A in a clockwise direction when viewed from the top. Where you can predict the magnitude of the magnetic field inside the loop will be maximum by taking symmetry into consideration? •
12. Could Figure P28.12 represent the magnetic field due to a bar magnet that has a rectangular cross section rather than a circular cross section? •

Figure P28.12



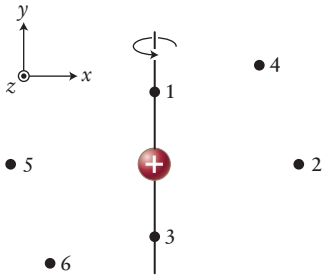
13. A loop of wire carries a current I as shown in Figure P28.13. Determine the direction of the magnetic field at the points labeled. •

Figure P28.13



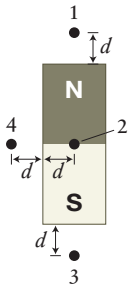
14. Figure P28.14 shows a positively charged spherical object that is spinning. For each of the labeled points, state the direction of the magnetic field. Assume all points lie in the xy plane. ••

Figure P28.14



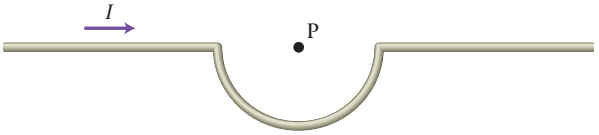
15. Consider four identical square current loops arranged in a square with four times the area of one loop. Use this arrangement to predict how the magnetic field far from a small current loop depends on the area of the loop if the current and orientation of the loop is held fixed. ••
16. Consider two charged particles 1 and 2, each of them able to translate (move from place to place) and to spin. In which of the following circumstances is there a magnetic interaction between the particles: (*a*) neither particle translating, neither particle spinning; (*b*) neither translating, only particle 1 spinning; (*c*) neither translating, both spinning; (*d*) 1 at rest and not spinning, 2 translating and not spinning; (*e*) 1 at rest and not spinning, 2 translating and spinning; (*f*) 1 spinning and not translating, 2 translating and not spinning; (*g*) 1 spinning and not translating, 2 translating and spinning; (*h*) both translating, neither spinning; (*i*) both translating, only 2 spinning; (*j*) both translating, both spinning. ••
17. A circular current-carrying wire loop produces a magnetic field. A spinning disk of uniformly distributed charge produces a similar magnetic field. How would you expect the field due to a spinning disk to compare to the field due to a current-carrying wire loop? ••
18. For the bar magnet in Figure P28.18, which of the four locations labeled 1, 2, 3, 4 has the greatest density of magnetic field lines? ••

Figure P28.18



19. A negatively charged particle located at the origin of an xyz coordinate system is spinning clockwise about the x axis (that is, clockwise when you look at the particle while standing somewhere on the positive x axis). What is the direction of its magnetic field at all positions along (*a*) the x axis, (*b*) the y axis, and (*c*) the z axis? ••
20. Figure P28.20 shows a wire segment bent into a half-circle, with the center of curvature labeled P. If the wire of which this segment is a part is extremely long and carries a current I, what is the direction of the magnetic field at P? ••

Figure P28.20

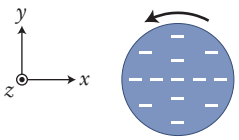


21. In the Bohr model of the hydrogen atom, an electron orbits a nucleus consisting of one proton. Given that the electron and proton are both spinning, describe the types of magnetic interactions you expect to be associated with the hydrogen atom. •••

28.3 Magnetic dipole moment and torque

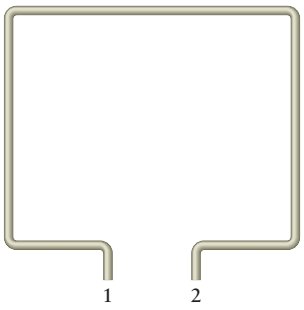
22. In the space surrounding Earth, the main component of the planet's magnetic field is the field due to a magnetic dipole. In what direction is Earth's magnetic dipole moment? •
23. Earth's magnetic field is thought to be generated by currents in the planet's core. In what direction (clockwise or counterclockwise, when viewed along the rotation axis from north to south) does this current circulate? •
24. The disk shown in Figure P28.24 is charged with electrons and spins counterclockwise when viewed from above. From this perspective, what is the direction of the magnetic dipole moment of the disk? •

Figure P28.24



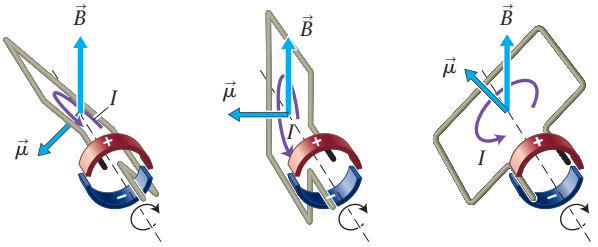
25. In the open wire loop of Figure P28.25, end 1 is held at a higher electric potential than end 2. (*a*) In which direction do electrons move through the loop? (*b*) What is the direction of the current? (*c*) What is the direction of the loop's magnetic dipole moment? ●●

Figure P28.25



26. Figure P28.26 shows part of the rotation of an electric motor, during which the magnetic dipole moment of the loop points to the left or has some component to the left. If the entire rotation were shown, during what fraction of the rotation would the magnetic dipole moment have some component pointing to the right? ●●

Figure P28.26

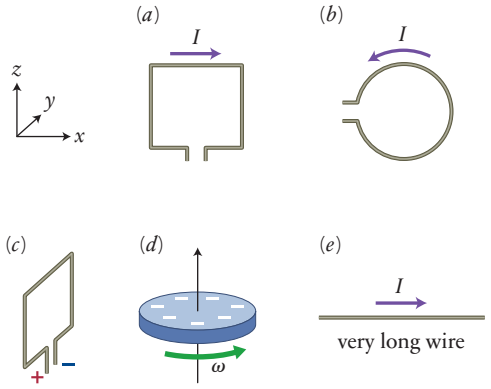


27. A current loop lying in the *xy* plane of an *xyz* coordinate system experiences a torque that is clockwise about the *y* axis when viewed looking down from the positive *y* axis toward the origin. This torque is due to a uniform external magnetic field that points in the positive *x* direction. What is the direction of the current in the loop? ●●

28. A circular current loop lies in the *xy* plane of an *xyz* coordinate system and is initially oriented so that it carries a clockwise current when viewed from the positive *z* axis. It is mounted in such a way that it can rotate about the *x*-axis, which passes through its centre, without affecting the magnitude or direction of the current it carries. How would the loop rotate if the magnetic field is rotated by a small angle about the *x*-axis? ●●

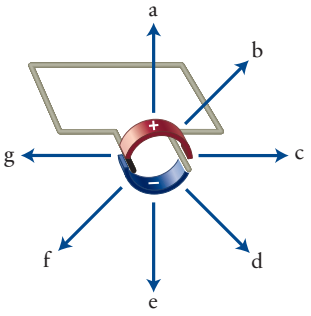
29. Determine the direction of the magnetic dipole moment in each current loop or charge distribution in Figure P28.29. In (*c*), the higher potential end of the loop is marked +, and in (*d*), the disk has a uniform negative charge. ●●

Figure P28.29



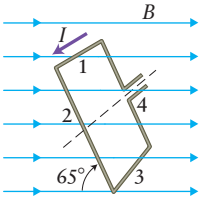
30. (*a*) As the electric motor shown schematically in Figure P28.30 operates, which of the arrows shown could represent the magnetic dipole moment at various instants? (*b*) If there are any arrows that are not possible magnetic dipole moment vectors, state why they are not allowed. ●●

Figure P28.30



31. Figure P28.31 shows a rectangular loop of current in an external magnetic field. Initially the plane of the loop makes a 65° angle with the magnetic field. (*a*) Determine the direction of the magnetic force exerted on each side of the loop. Treat side 4 as though it were one unbroken wire. (*b*) Determine the direction of the torque on the loop about the axis shown as a dashed line. (*c*) What is the orientation of the loop once the system reaches equilibrium? ●●

Figure P28.31



32. Figure 28.10 shows a current loop in an external magnetic field and the forces exerted on the different lengths of the wire. Assume that the loop is attached to a pivot

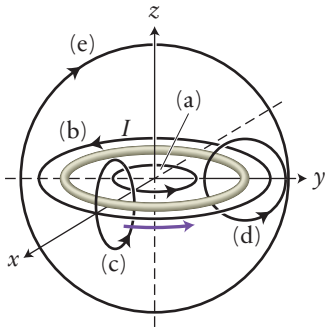
that allows it to spin (as between Figures 28.10 and 28.11). (*a*) Determine the directions of all the forces on the loop if the current were reversed. (*b*) Is this new situation any different from that shown in Figure 28.11? •••

33. If a stationary electron experiences a torque when a proton passes near it, then by the principle of relativity, a moving electron also experiences a torque when passing by a stationary proton. The electron orbiting the proton in a hydrogen atom, therefore, experiences a torque due to its intrinsic magnetic dipole moment. If the electron orbits counterclockwise in the *xy* plane when viewed from the positive *z* axis, what is the direction with which this torque tends to align the electron's intrinsic magnetic dipole moment? •••

28.4 Ampèrian paths

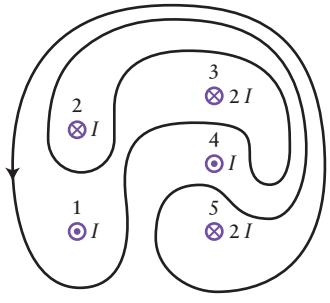
34. The current loop in Figure P28.34 lies in the *xy* plane. For each of the Ampèrian paths (a)–(e), is the line integral of the magnetic field positive, negative, or zero? •

Figure P28.34



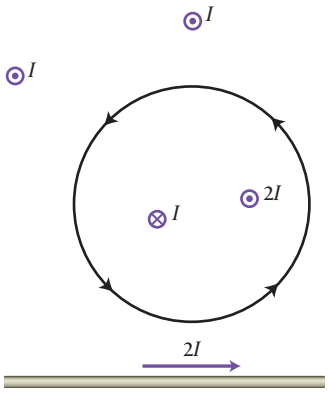
35. Wires 1 to 5 in Figure P28.35 carry current either into or out of the page. What is the magnitude of the current enclosed by the Ampèrian path indicated? Is the line integral of the magnetic field around this path, in the direction indicated, positive, negative, or zero? •

Figure P28.35



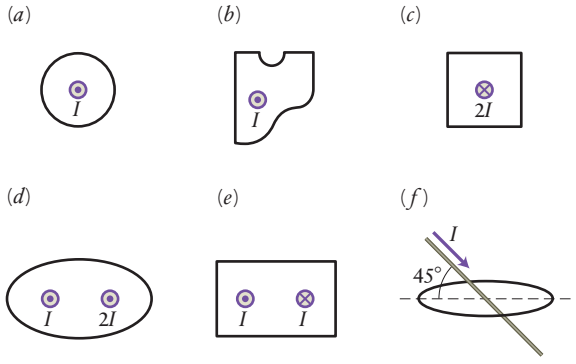
36. Is the line integral of the magnetic field along the closed path in Figure P28.36 positive, negative, or zero? The direction of integration is shown with arrows on the path. •

Figure P28.36



37. Figure P28.37 shows a series of current-carrying wires, and in each case an Ampèrian path is shown (but without direction). Rank the six cases according to the magnitude of the line integral of the magnetic field calculated along the Ampèrian path shown in each case, smallest value of $\left|\oint \vec{B} \cdot d\vec{\ell}\right|$ first. ••

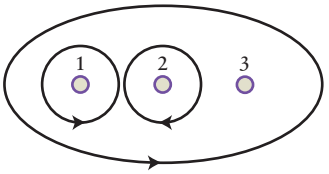
Figure P28.37



38. The line integral of the magnetic field around a certain closed path is initially *L*. The current that penetrates this path is then doubled. What can you say about the current through the path if the new integral of the magnetic field around the closed path is (*a*) *L* and (*b*) 2*L*? ••

39. Each of the wires 1 to 3 in Figure P28.39 carries a current perpendicular to the page. The line integrals of the magnetic field around the three Ampèrian paths shown all have the same positive value. How do the magnitudes and directions of the three currents compare? ••

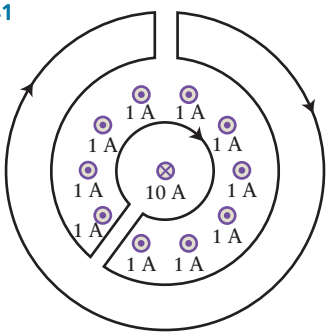
Figure P28.39



40. A positively charged particle located at the origin of an *xyz* coordinate system spins about the *z* axis, and the spin is counterclockwise when viewed looking down from the positive *z* axis toward the origin. Is the line integral of the magnetic field positive, negative, or zero for an Ampèrian path going (*a*) clockwise around the origin in the *xy* plane when viewed looking down from the positive *z* axis, (*b*) clockwise around the origin in the *yz* plane when viewed looking from the positive *x* axis, and (*c*) clockwise around the origin in the *xz* plane when viewed looking from the positive *y* axis? ●●

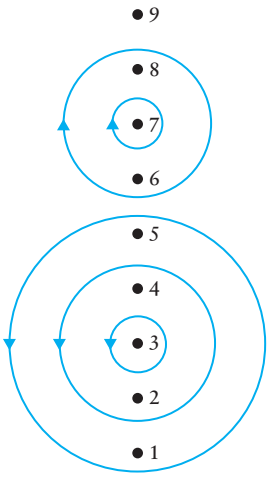
41. Eleven wires and one Ampèrian path are shown in Figure P28.41, with current values and directions as indicated. Is the line integral of the magnetic field along the Ampèrian path shown greater than, smaller than, or equal to zero? ●●

Figure P28.41



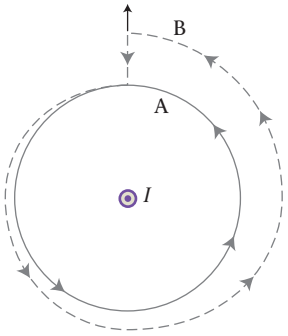
42. Figure P28.42 shows magnetic field lines in a certain region. Are there any locations where there must be a current directed into or out of the plane of the diagram? (Consider locations near the numbers on the vertical axis.) At which of the locations 1–9 does a current exist? What is the current direction in each case? ●●

Figure P28.42



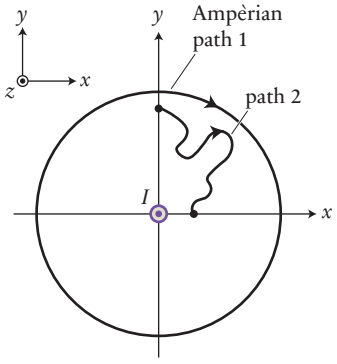
43. Figure P28.43 shows two paths (A and B) around a wire that carries current *I*. (*a*) Along which path is the line integral of the magnetic field greater? (*b*) Along which path is the average magnetic field greater? (*c*) Explain how your answers to parts *a* and *b* are consistent. ●●●

Figure P28.43



44. Figure P28.44 shows a very long wire that carries a current *I* in the *z* direction and is centered in the *xy* plane at the position (0, 0). The line integral of the magnetic field along Ampèrian path 1 is $-5.50\ \mathrm{T \cdot m}$. Determine the line integral of the magnetic field along path 2. ●●●

Figure P28.44



28.5 Ampère's law

45. A long, straight, current-carrying wire carries a current of magnitude 3 A. Calculate the magnitude of the magnetic field at a location 35 mm away from the wire. ●

46. A long, straight wire carrying a current of 5 A flowing from left to right is placed above a large, flat sheet through which the current per unit width is 6.0 A/m flows from right to left. Calculate the magnitude and direction of the magnetic force per unit length exerted by the wire on the sheet. ●

47. The magnitude of magnetic field around a straight current carrying wire at a radial point 25 mm is 6 mT. (*a*) At what radial point will the magnitude of magnetic field be 10 mT? (*b*) Calculate the current in the wire. ●

48. Two infinitely long, straight, parallel wires are 6 cm apart, carrying currents 2 A and 4 A, respectively, in the same direction. A third infinitely long, straight wire, parallel to the other two, and carrying a current *I*, lies between the two wires. Find the equilibrium position(s) of this wire. ●●

49. An electron travels at a velocity of 3.0×10^6 m/s parallel to a sheet carrying a current to the right. When a uniform electric field of magnitude 200 V/m is applied, it travels at an angle of 45° to the current in the sheet, with no deflection. What is the magnitude of the current per unit width in the sheet? ●●

50. An electric current is uniformly distributed throughout a long, straight wire that has a diameter of 10 mm. If the current through the wire is 3.0 A, calculate the magnitude of the magnetic field (*a*) 4 mm radially away from the wire center and (*b*) 12 mm radially away from the wire center. (*c*) What must the current be to create a magnetic field of magnitude 10^{-2} T at a distance of 5 cm from the wire center? ••

51. Two large, flat current-carrying sheets are placed parallel to each other, one sheet above the other. The upper sheet carries a current 1.0 A per unit of width to the left, and the lower sheet carries a current 3.0 A per unit of width to the right. Calculate the magnitude of the magnetic field (*a*) between the sheets, (*b*) above the upper sheet, and (*c*) below the lower sheet. (*d*) Draw a diagram showing the direction of the magnetic field in each of these regions. ••

52. An electron travels to the right at 2.5×10^6 m/s between two large, flat sheets that are parallel to each other and to the electron's line of motion. A current per unit width of 8.0 A/m flows to the right through the top sheet. (*a*) What current should flow in the second sheet so that the electron does not deviate from its path? (*b*) If one generates a current of 10 A/m to the right on the second sheet, what is the magnitude and direction of the magnetic force on the electron? ••

53. A certain wire has a circular cross section of radius R and carries a current I. Suppose that the charge carriers all move along the cylindrical surface of the wire, not through its cross-sectional area. (*a*) Derive an expression for the magnetic field magnitude $B(r)$ as a function of distance r from the center of the wire; check that your expression makes sense for $r < R$ and for $r > R$. (*b*) Make a graph showing the magnitude of the magnetic field in and around the wire as a function of the radial distance r from the center. Mark the wire radius R on your graph. ••

54. An electron travels at a constant velocity at a distance $d_1 = 4.0$ mm above a large sheet of metal that carries a current of 3 A, and at a distance $d_2 = 3.0$ mm below a very long wire that carries a current 0.5 A. The wire is parallel to the sheet, and both currents flow in the same direction. Compute the width of the metal sheet. ••

55. Two large, parallel, current-carrying plates are oriented horizontally and the vertical distance between them is 5.0 mm. The current per unit width in each plate is 100 A/m, and both currents are in the positive x direction. Determine the magnitude and direction of the magnetic force per unit area exerted by the lower plate on the upper plate. ••

56. A coaxial cable carries a current I_{wire} in its inner conducting wire and a current I_{shell} in its outer conducting shell. The radius of the wire is R_{wire}, the distance from the cable center to the inside of the shell is R_{shell}, and the distance from the center to the outer edge of the shell is $2R_{\text{shell}}$. Determine all locations (if any exist) at which the magnetic field magnitude is zero if (*a*) $I_{\text{wire}} = I_{\text{shell}}$, (*b*) $I_{\text{wire}} = -I_{\text{shell}}$, and (*c*) $I_{\text{wire}} = -I_{\text{shell}}/2$. •••

57. A particle is shot at a velocity of 3×10^4 m/s to the right and enters the magnetic field generated by a large, flat current-carrying sheet. The current in the sheet is parallel to the particle's initial line of motion. The particle exits the magnetic field 6.8 cm directly above the location where it enters. (*a*) Is the particle's charge positive or negative? (*b*) If the sheet carries a current of 4 A per unit width, what is the ratio of the charge on the particle to its mass? •••

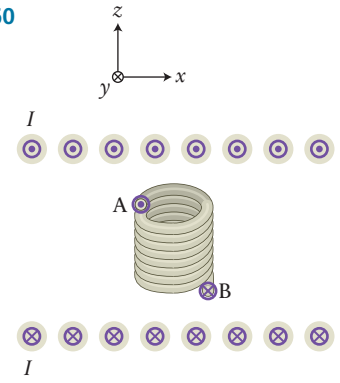
28.6 Solenoids and toroids

58. A long solenoid having 16 windings per centimeter of length carries a current of 0.8 A. Calculate the magnitude of magnetic field inside the solenoid. •

59. Considering the magnetic field produced at the center of solenoid is uniform, one need to produce a magnetic field of magnitude 0.06 T. If the available current that one can use to flow through the windings is 12 A, calculate the number of windings per centimeter the solenoid must have. •

60. A small solenoid is inserted into a larger solenoid (Figure P28.60). The current in the small solenoid is from A to B. (*a*) Determine the initial direction of the magnetic dipole moment of the small solenoid. (*b*) Assuming the large solenoid is fixed in place and the small solenoid is free to rotate, determine the equilibrium direction of the magnetic dipole moment of the small solenoid. •

Figure P28.60



61. A long solenoid with 1000 windings per meter is placed with its central axis parallel to the y axis and 20 mm above a sheet of metal that is 0.5 m wide in the y direction and carries a uniform current of 20 A in the positive x direction. What current must the solenoid carry in order that the magnetic field at its centre is zero? ••

62. A small coil, of area 2 cm^2, carries a current of 1.5 A. The coil is suspended inside a solenoid at an angle of 45° with respect to the magnetic field. The solenoid has 500 turns per meter of length and carries a current of 4.0 A. Calculate the magnitude of the torque exerted on the coil. ••

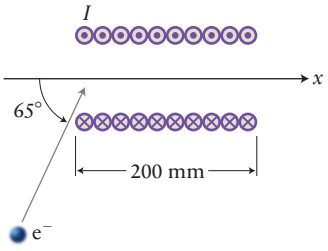
63. A toroid has 250 square windings carrying a current of 3.0 mA. Each side of each square winding is 50 mm long, and the distance from the toroid center to the inner surface of the windings is 120 mm. What is the magnitude of the magnetic field (*a*) at the center of any square winding, (*b*) midway between the toroid center and the inner surface of the windings, and (*c*) 30 mm beyond the outer surface of the windings? ••

64. If the magnitude of magnetic field produced at a radial distance of 8.5×10^{-2} m from the center of a toroid of 300 windings is equal to the magnetic field produced inside a solenoid, then calculate the number of turns per unit length of the solenoid. Consider the same current flowing through both the solenoid and toroid. ••

65. A current of 10 A flows through a toroid that has 300 circular windings per meter measured along the interior edge of the windings. The radius of each circular winding is 20 cm, and the inner radius of the toroid is 40 cm. Find the magnitude of the magnetic field at 30 cm inside the circular winding. ••

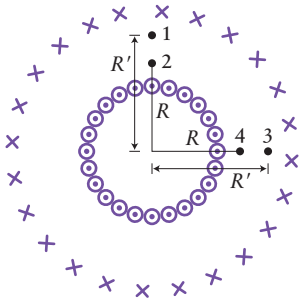
66. An electron is fired into one end of the solenoid in Figure P28.66. Viewed along the positive x axis from a negative x coordinate, the electron enters from below at a 65° angle to the horizontal, just inside the bottom edge of the solenoid. From this viewpoint the solenoid carries a 10-A clockwise current. The solenoid is made from a 33.0-m length of wire, and it has 400 turns along the 200-mm length shown in the figure. (*a*) Ignoring end effects, what is the smallest time interval required for the electron to pass through the solenoid without striking the coils? (*b*) If the electron follows the quickest path through the solenoid as determined from the time interval you calculated in part *a*, how many revolutions does the electron's path make around the solenoid axis? •••

Figure P28.66



67. You are calibrating magnetic coils for a particle detector. One step involves checking the magnetic field at different positions inside a toroid, and you are asked to measure the field at the positions labeled 1, 2, 3, 4 in Figure P28.67. Your partner, knowing that the magnetic field inside a *solenoid* is uniform, asks why anyone would measure more than one position inside the toroid, which is a solenoid bent into a ring. How can you explain the need to compare the field magnitudes at positions 1 and 2? Why compare the magnitudes at positions 3 and 4? Can you explain to your partner why the magnetic field is not uniform in a toroid even though a toroid is nothing more than a solenoid bent into a donut shape? •••

Figure P28.67



28.7 Magnetic fields due to currents

68. Calculate the magnitude of the magnetic field at the center of a circular arc of radius 25 mm spanning an angle of $\pi/2$ and carrying a current of 3.0 A. •

69. A 100-mm section of wire is bent into a circular arc. What must the radius of the arc be for a current of 6.0 A to produce a magnetic field of 7.5×10^{-5} T at the centre of the arc? •

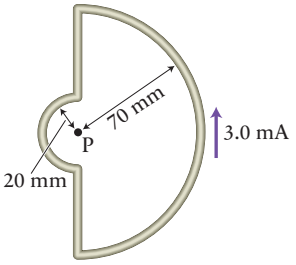
70. Use the Biot-Savart law to determine the magnetic field 70 mm above the center of a loop of wire that has a radius of 0.22 m and carries 3.0 A of current. ••

71. Wire 1, 3.0 m long and of linear mass density 0.010 kg/m, is initially held in place and carries 10 A of current to the right. Very long wire 2 is placed parallel to and 5.0 mm directly above wire 1. When wire 1 is released, it stays in place. What are the magnitude and direction of the current in wire 2? ••

72. The magnetic field a distance z above the centre of a circular loop of radius r carrying current I in the xy plane is $\vec{B} = \frac{\mu_0 I r^2}{2(r^2+z^2)^{3/2}}\hat{k}$. [See problem 28.70] (*a*) Use this result to re-derive the formula for the magnetic field inside a very long solenoid of radius r with n windings per unit length, carrying a current I per winding. [Hint: you need to perform an identical integral to the one required to calculate the magnetic field of a straight wire using the Biot-Savart law.] (*b*) Now replace the very long solenoid with one of arbitrary length L, and find the magnetic field on its central axis, a distance D from its centre. ••

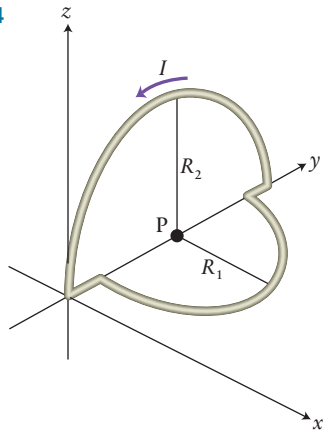
73. In Figure P28.73, point P is the common center of two circular arcs of wire, the larger of radius 70 mm and the smaller of radius 20 mm. What are the magnitude and direction of the magnetic field at P if the current through the wire is 3.0 mA? ••

Figure P28.73



74. A current-carrying wire has been bent into the form shown in Figure P28.74, with a half-circle of radius R_1 lying in the xy plane connected via two straight segments to a half-circle of radius R_2 lying in the yz plane. The current direction is as indicated. If $R_2 = 1.5R_1$, what is the direction of the magnetic field at P, which lies at the common center of the half-circles? ••

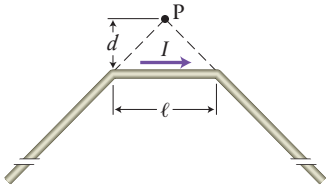
Figure P28.74



75. (*a*) Find the magnetic field produced by a 10-cm section of a wire carrying a 3 A current at a distance 5 cm above its centre. (*b*) Now, find the magnetic field at the centre of a square loop comprising four such sections. •••

76. The horizontal portion of the wire in Figure P28.76 has a length $\ell = 0.100$ m, and position P is a perpendicular distance $d = 30.0$ mm above the center of the horizontal portion. The wire carries a 45.0-A current, and the two slanted portions of the wire are extremely long. Calculate the magnitude of the magnetic field at P. •••

Figure P28.76



28.8 Magnetic field of a moving charged particle

77. An electron moves in a straight line at a speed of 6.0×10^7 m/s. Calculate the magnitude and direction of the magnetic field at a position 5.0 mm behind the electron and 15 mm below its line of motion. •

78. A proton moves in the positive x direction at 4.00×10^4 m/s. Calculate the magnitude of the magnetic field at the point $(x, y) = (+2.00 \text{ mm}, +1.00 \text{ mm})$ as the proton passes through the origin. •

79. Express the magnetic field due to a uniformly moving charged particle in terms of the electric field of the particle and its velocity. ••

80. Two electrons 1 and 2 move along antiparallel paths separated by a distance of 10 nm, traveling at speeds $v_1 = 4.0 \times 10^7$ m/s and $v_2 = 7.0 \times 10^6$ m/s. What is the magnitude of the magnetic force exerted by electron 2 on electron 1? ••

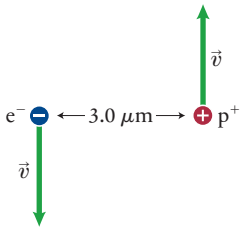
81. Proton 1, traveling in the negative x direction at speed v_1, is directly below proton 2, which is traveling at speed v_2 at an angle of 45° above the positive x axis. If the protons are separated by a distance r, calculate the magnitude of the magnetic force each proton exerts on the other. ••

82. One runs into trouble when trying to apply Ampère's law to the magnetic field of a moving charge. For instance, what is the current enclosed by an Ampèrian path that encircles the path of the particle? Because of the discrete nature of the charge, any sensible denition is zero most of the time and very large for a brief instant. The magnetic field does not have a similarly abrupt time dependence. One way out is to integrate over time. Show that the magnetic field $\vec{B}$ produced by a charge q moving in a straight line at a constant velocity obeys $\int_{-\infty}^{\infty} \left(\oint \vec{B} \cdot d\vec{\ell} \right) dt = \mu_0 q$. Use a circular Ampèrian path in the plane perpendicular to and centred on the path of the charge. ••

83. The electromagnetic force between two particles, both of mass m, with one having charge $+q$ and the other having charge $-q$ can cause them to orbit in a circle around the centre of the line that connects them. (*a*) Relate their velocities to their separation distance d. (*b*) Without the magnetic contribution to the force, would the velocity at a given separation be larger or smaller? ••

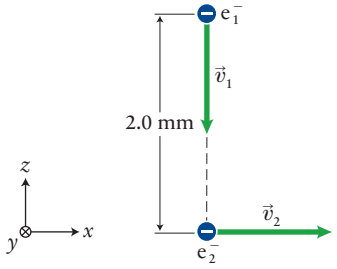
84. An electron and a proton are fired in opposite directions, and at the instant they are nearest each other, their separation distance is 3.0 μm (Figure P28.84). At that instant, the two particles are both moving at 3.0×10^4 m/s but in opposite directions. (*a*) Determine the magnitude and direction of the magnetic force exerted by the proton on the electron at this instant. (*b*) Show that the ratio of the magnetic to electric forces at this instant equals v^2/c_0^2, where $v = 3.0 \times 10^4$ m/s and c_0 is the speed of light. ••

Figure P28.84



85. Two electrons move near each other and at the instant shown in Figure P28.85 are 2.0 mm apart. The speed of electron 1 is $v_1 = 300$ m/s, that of electron 2 is $v_2 = 500$ m/s, and the directions of motion are as shown in Figure P28.85. Determine the magnitude and direction of (*a*) the magnetic force exerted by electron 1 on electron 2 and (*b*) the magnetic force exerted by electron 2 on electron 1. ••

Figure P28.85



86. At $t = 0$, electron 1 is shot out of an accelerator at a speed of 2.0×10^3 m/s. At $t = 1.0$ μs, electron 2 is shot out of the accelerator and travels on a path that is parallel to and 10 mm below the path of electron 1. The speed of electron 2 is 5.0×10^3 m/s. (*a*) At $t = 3.0$ μs, what is the magnitude of the magnetic field at a point that is 3.0 mm ahead of electron 1 and 5.0 mm below it? Assume that the magnetic forces the electrons exert on each other for the 2.0 μs before you make your calculations do not alter the electrons' paths significantly. (*b*) Show whether or not it is valid to assume that the magnetic forces the electrons exert on each other during the interval from $t = 1.0$ μs to $t = 3.0$ μs do not alter the electrons' paths significantly. ●●●

Additional Problems

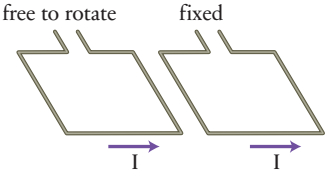
87. A current loop lies in the *xy* plane, with the current circulating counterclockwise when viewed from the positive *z* axis. Is there a torque on the loop, and if so, in what direction, if a uniform magnetic field is applied along the positive (*a*) *x* axis, (*b*) *y* axis, and (*c*) *z* axis? ●

88. The length of a straight solenoid is 50 cm. If a current of 12 A flows through the solenoid produces a magnetic field of magnitude 0.06 T inside it, calculate the number of windings. ●

89. Two parallel rods having length 12 cm and a separation of 50 cm carry current from left to right. If current through wire 1 is 10 A and through wire 2 is 12 A, calculate the magnitude and direction of the magnetic force per unit length exerted by wire 1 on wire 2. ●●

90. Figure P28.90 shows two square loops of wire. The loop on the right is fixed in place, and the one on the left is free to pivot in any direction. Both loops carry a current that is counterclockwise when viewed from above. (*a*) What is the direction of the magnetic field created by the right loop at the left loop? (*b*) What is the initial direction of the magnetic dipole moment of the left loop? (*c*) How does the left loop rotate? (*d*) After the left loop is allowed to rotate and the system reaches equilibrium, are the magnetic dipole moments of the two loops parallel, antiparallel, or neither? ●●

Figure P28.90



91. A wire segment of 1.0 m and carrying a current of 5.0 A is held in a plane in which it can rotate and translate with little friction. A long wire is fixed parallel to this plane, perpendicular to the 1.0 m wire and 50 mm above its centre. (*a*) Describe how the 1.0 m wire starts to move when a 10-A current is switched on in the long wire. (*b*) Calculate the magnetic force on the 1.0 m wire after static equilibrium is reached. ●●

92. You have a solenoid for which a current *I* per winding produces a magnetic field of magnitude B_1 inside the soleniod. You cut the solenoid open parallel to its symmetry axis. You then carefully bend it open to form a flat sheet of parallel wires and electrically connect the wire ends on the cut edges in a way that allows you to run the same current through every parallel wire. Now a current *I* per wire produces a magnetic field of magnitude B_2 just above the sheet. You have a friend who claims that B_1 must be many times bigger than B_2, because, by cutting the solenoid open, you have allowed the field lines trapped inside the solenoid to escape into an infinite volume. You have another friend who claims that $B_1 = B_2$, because in both setups, the magnetic field can be determined from Ampère's law using the same rectangular path that encloses the same current. Explain why both friends are wrong. ●●● CR

93. You need to design a square toroid such that the magnetic field in the centre of the square windings is 0.0020 T, and the side length of a square winding is 3 cm. The wire you are using can carry a current of at most 0.5 A. The magnetic field anywhere inside the windings must be at least 0.0016 T. What is the minimum number of windings you will need? ●●● CR



29

Changing Magnetic Fields

29.1 Moving conductors in magnetic fields
29.2 Faraday's law
29.3 Electric fields accompany changing magnetic fields
29.4 Lenz's law

29.5 Induced emf
29.6 Electric field accompanying a changing magnetic field
29.7 Inductance
29.8 Magnetic energy

CONCEPTS

QUANTITATIVE TOOLS

Up to now, the only electric fields we have encountered are those that arise from the presence of electrical charges. In this chapter we explore electric fields that accompany changing magnetic fields. As we saw in Chapter 27, constant magnetic fields exert forces on moving charge carriers. Now we discover that there is no fundamental difference between the interaction of changing magnetic fields with stationary charge carriers and the interaction of constant magnetic fields with moving charge carriers. We shall also learn that energy can be stored in magnetic fields. These ideas are harnessed in a wide variety of important applications, from electric motors and electrical power generation to electronic appliances and other everyday devices.

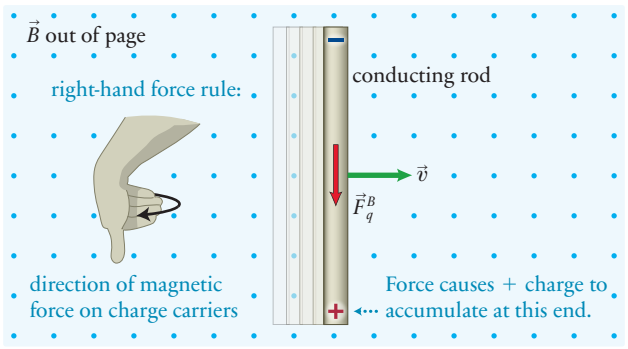
29.1 Moving conductors in magnetic fields

Let's examine what happens when a conducting rod moves with velocity $\vec{v}$ through a magnetic field of constant magnitude B, where $\vec{v}$ is perpendicular to $\vec{B}$ (**Figure 29.1**). Recall from Chapter 27 that a magnetic field exerts a magnetic force on moving charge carriers. That force is proportional to the field magnitude, the amount of charge, and the velocity component perpendicular to the field. The right-hand force rule gives the direction of the resulting force exerted on positive charge carriers.

As the rod in Figure 29.1 moves to the right in the magnetic field, positive charge carriers in the rod experience a downward force. As a result, a positive charge accumulates at the lower end of the rod, leaving behind a negative charge at the other end. (If the charge carriers are negatively charged, they experience an upward force; the end result, however, is still as shown in Figure 29.1.)

Does charge continue to accumulate as the rod moves? No—as charge accumulates, the electric field between the oppositely charged ends of the rod resists further accumulation of charge. Once the opposing force due to the electric field is equal in magnitude to the force exerted by the magnetic field on the charge carriers, the vector sum of the forces exerted on the charge carriers is zero, and no further accumulation of charge takes place. The motion of the rod through the magnetic field thus establishes a charge separation, the amount of which is determined by the magnitude of the force exerted by the magnetic field on the charge carriers.

Figure 29.1 When a conducting rod moves in a magnetic field, positive charge carriers in the rod experience a magnetic force. In the configuration shown, the right-hand force rule tells us that a positive charge accumulates at the bottom of the rod.



CONCEPTS

29.1 What happens if the rod in Figure 29.1 moves to the left?

Now consider what happens if, instead of a conducting rod, a conducting rectangular loop moves into a magnetic field (**Figure 29.2**). As the right side of the loop enters the field, its charge carriers move in response to the magnetic force exerted on them. Rather than just accumulating at the ends of the right side, however, the carriers now have a closed path along which to travel. Thus, the magnetic force exerted on the carriers creates a current in the loop. (The carriers in the top and bottom sides of the loop also feel magnetic forces, as shown in Figure 29.2, but because these carriers can't move very far in the vertical direction, we can ignore the effects of these forces.)

The current that arises from moving a conducting loop through a magnetic field can be verified experimentally. **Figure 29.3** shows a simple experiment using a bar magnet and a current meter, which shows both the magnitude and the direction of the current through the wire. When the wire is at rest in the field of the magnet (Figure 29.3*a*), the meter indicates no current. Moving part of the wire to the right causes a counterclockwise current (Figure 29.3*b*); moving the same part of the wire to the left causes a clockwise current (Figure 29.3*c*).

The current that arises from the motion of charged particles relative to a magnetic field is called an **induced current** (as opposed to a "regular" current caused by a potential difference between, for instance, the terminals of a battery).

Figure 29.2 Direction of current in a rectangular loop moving into a magnetic field. Magnetic forces exerted on the charge carriers in the loop cause them to move around the loop.

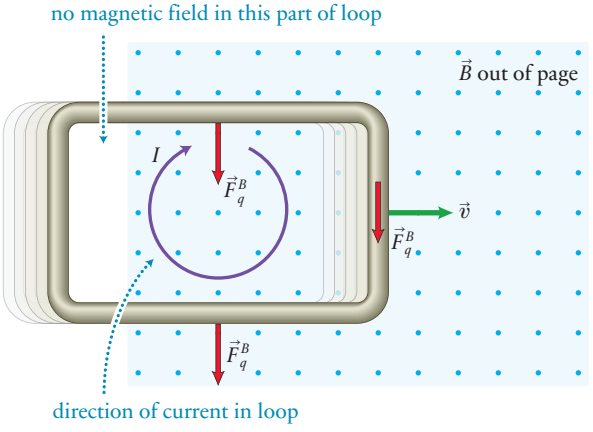
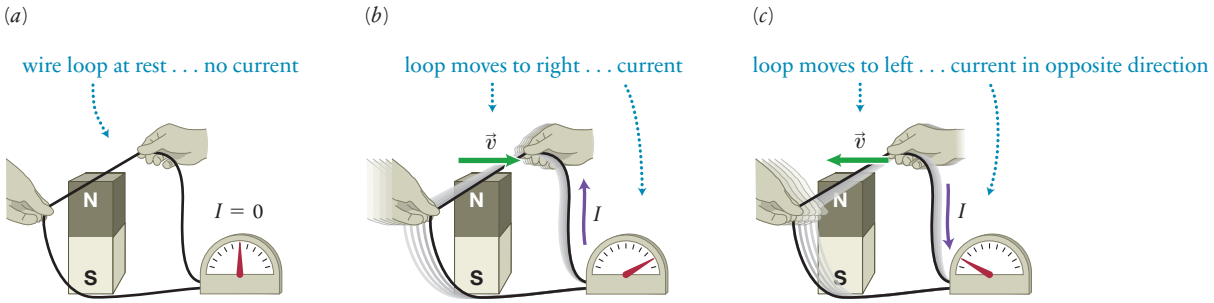


Figure 29.3 Experimental observation of induced current.



EXAMPLE 29.1 A rectangular loop passing through a magnetic field

Consider a rectangular conducting loop traveling from left to right through a magnetic field as shown in **Figure 29.4**. During this motion, at what positions *a*–*e* is a current induced in the loop, and what is the current direction at each position?

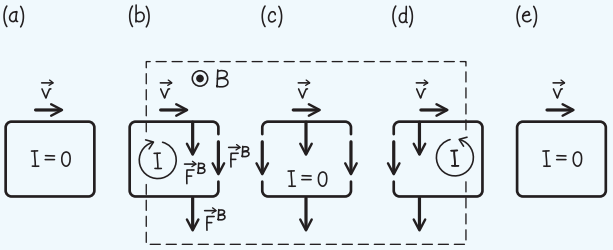
1 GETTING STARTED As the loop moves through the magnetic field, the charge carriers in the loop experience a magnetic force. For each of the five positions I need to determine whether or not the magnetic force causes the carriers to flow around the loop. To simplify the problem, I assume that only the positive charge carriers are free to move.

2 DEVISE PLAN The direction of the magnetic force exerted on the charge carriers is given by the right-hand force rule. To see if there is a current in the loop, I'll sketch the loop in each position and determine the direction of the magnetic force exerted on the charge carriers in each side of the loop.

3 EXECUTE PLAN No matter where along the loop perimeter the carriers are, they always move to the right with the loop, and so curling the fingers of my right hand from $\vec{v}$ to $\vec{B}$, I see that the magnetic force exerted on them is always toward the bottom of the page. I draw this force for each side of the loop that is inside the magnetic field (**Figure 29.5**). The magnetic forces exerted on the top and bottom sides of the loop cause no current through the loop because the charge carriers in these sides have essentially no mobility in the direction in which $\vec{F}^B$ is exerted.

In position (*a*), with no part of the loop in the magnetic field, no magnetic force is exerted on any of the carriers, and therefore no current is induced. When the right side of the loop passes into the magnetic field but the left edge has not yet entered (*b*), the magnetic force exerted on the charge carriers in the right side of the loop causes a clockwise current around the loop. Once the entire loop is in the magnetic field (*c*), the magnetic force exerted on the carriers in the left side drives a counterclockwise current equal in magnitude to the clockwise current driven by the magnetic force exerted on the carriers in the right side. The combined effect is that no current is induced. In position (*d*), only the left side of the loop is in the field, and the magnetic force exerted on those charge carriers induces a counterclockwise current around the loop. When the loop has passed out of the magnetic field (*e*), once again no magnetic forces are exerted on the carriers and therefore no current is induced. ✔

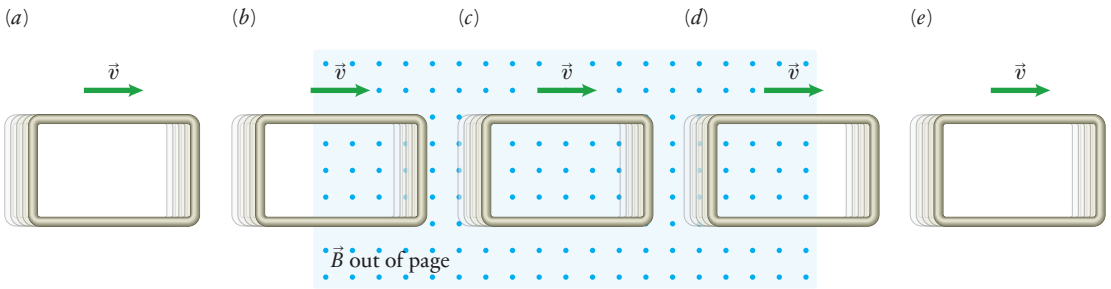
Figure 29.5



4 EVALUATE RESULT In position (*b*) I have the same situation as in Figure 29.2, so it is reassuring that my answer is the same. If I ignore any dissipation of energy, the initial and final states are the same. Therefore it makes sense that the effect of moving the loop through the field is zero: The charge carriers first move clockwise as the loop moves into the field, then move counterclockwise as the loop moves out of the field.

(Continued)

Figure 29.4 Example 29.1.



CONCEPTS

I assumed that only the positive charge carriers can move. Had I assumed that the negative charge carriers are free to move, my answer would not have changed. Even though the direction of motion of the negative charge carriers is opposite the direction of motion of the positive carriers, the current is still in the directions shown in Figure 29.5 because current direction is *defined* as the direction in which positive charge carriers move.

29.2 In Example 29.1, suppose the loop is stationary and the source of the magnetic field is moved to the left such that their relative motion is the same. Do you expect there to be a current through the loop?

29.2 Faraday's law

As Example 29.1 indicates, a current is induced through a conducting loop when the loop moves into or out of a uniform magnetic field, not when the loop is entirely in or entirely out of the field. In other words, a current is induced when the area of the loop inside the field changes—that is, when the magnetic flux passing through the loop is changing with time.

Experiments show that the rate at which the magnetic flux through the loop changes affects the magnitude of the induced current. As illustrated in Figure 29.6, moving the loop through the magnetic field faster, which causes a greater rate of change in the magnetic flux through the loop, produces a greater induced current.

As we learned in Chapters 6 and 14, physical phenomena do not depend on the reference frame in which we observe them. Indeed, if the change in magnetic flux through a conducting loop is responsible for inducing a current, it should not matter whether the loop is moving in a constant magnetic field or the source of the magnetic field is moving and the loop is stationary. Only the motion of one relative to the other matters. Experiments confirm that the same current is measured in the loop no matter whether the loop moves to the right with speed v or the source of the magnetic field moves to the left with speed v (Figure 29.7). In fact, current is induced in a stationary conducting loop in any region where a magnetic field is changing, even if the source of the magnetic field is stationary and the magnitude of the magnetic field is changing.

Figure 29.6 The speed at which a conducting loop is moved through a magnetic field affects the magnitude of the induced current.

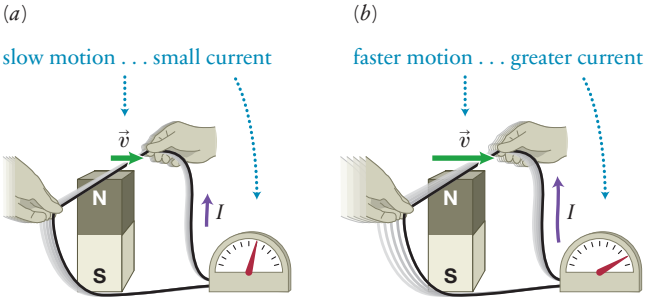
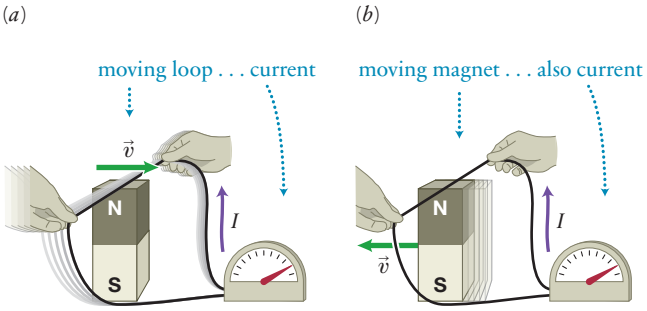


Figure 29.7 A current can be induced by moving either the loop or the magnet.



These experiments lead us to conclude:

A changing magnetic flux through a conducting loop induces a current in the loop.

This statement is called **Faraday's law.** The process by which a changing magnetic flux causes charge carriers to move, establishing a separation of charge or inducing a current, is called **electromagnetic induction.**

EXAMPLE 29.2 Changing magnetic flux

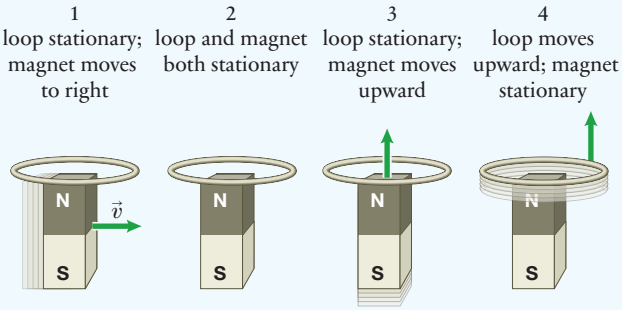
(*a*) In which of the four loops shown in Figure 29.8 is there a current that is caused by a magnetic force? (*b*) In which situation(s) is a current induced in the loop?

1 GETTING STARTED I begin by making a sketch of the magnetic field of the bar magnet, based on what I learned in Chapter 27 (Figure 27.13). I show a cross section of the loop placed above the north pole of the magnet (Figure 29.9).

2 DEVISE PLAN Magnetic forces are exerted on charge carriers moving relative to a magnetic field, provided there is a nonzero component of the magnetic field perpendicular to the carriers' velocity. To answer part *a*, therefore, I must determine which of the four situations satisfies those conditions. For part *b* I know that a changing magnetic flux induces a current, and so for each situation I must establish whether or not the magnetic flux through the loop is changing.

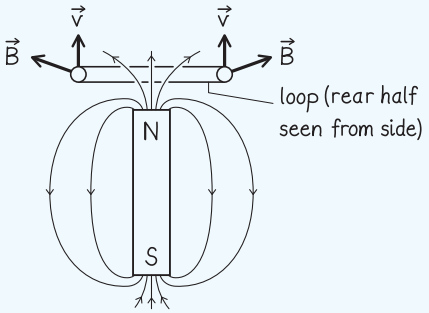
3 EXECUTE PLAN (*a*) Because the loop is stationary in situations 1 through 3, there cannot be a magnetic force in those

Figure 29.8 Example 29.2.



CONCEPTS

Figure 29.9



situations. For situation 4, I draw vectors in Figure 29.9 indicating the directions of the loop velocity and of the magnetic field at the position of the ring and the velocity of the loop. Because the magnetic field lines spread outward from the north pole, the magnetic field is not parallel to the velocity, and so a magnetic force is exerted on the charge carriers. According to the right-hand force rule, the force is in opposite directions on the two sides of the loop, so this magnetic force creates a current in the loop. ✔

(*b*) Because the field of a bar magnet is not uniform, any relative motion between the magnet and the loop causes the magnetic flux through the loop to change. Therefore, the magnetic flux is changing in situations 1, 3, and 4, which means current is induced in the loop in these three cases. ✔

4 **EVALUATE RESULT** That there is a current created by a magnetic force only in situation 4 makes sense because a magnetic force requires a nonzero velocity of the charge carriers and only in situation 4 is the loop moving. That there is an induced current in every situation except 2 makes sense because Faraday's law tells me that a current is induced whenever there is a changing magnetic flux through the loop—that is, whenever the loop and magnet move relative to each other.

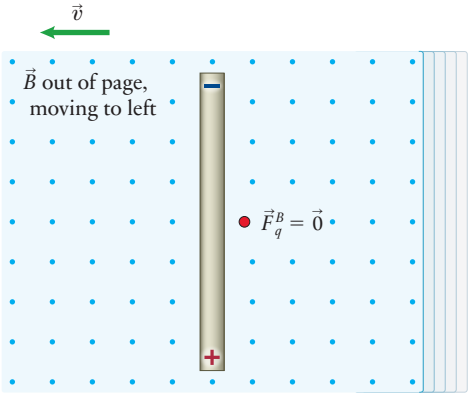
29.3 Is a magnetic force exerted on the (stationary) charge carriers in the loop of wire held above the magnet in Figure 29.7*b*?

29.3 Electric fields accompany changing magnetic fields

Example 29.2 and Checkpoint 29.3 lead to a surprising conclusion: Although no magnetic force is exerted on the charge carriers in a stationary loop, a current is still induced! **Figure 29.10** shows this situation in more detail. Experiments show that as the source of a magnetic field moves past a stationary conducting rod, a charge separation and hence a potential difference develop between the ends of the rod even though no magnetic force is exerted on stationary charge carriers.

The potential difference that develops between the ends of the rod shown in Figure 29.10 is the same as that which would develop if the source of the magnetic

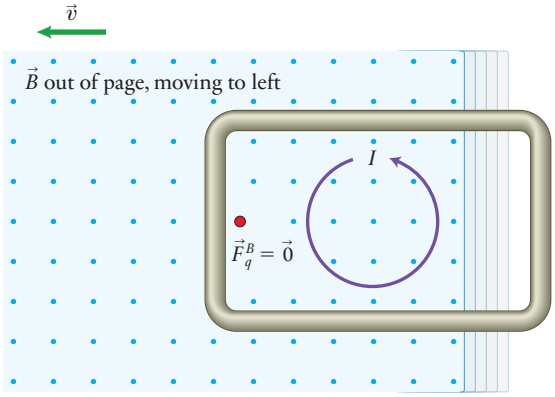
Figure 29.10 A stationary conducting rod in a moving magnetic field develops a charge separation, but no magnetic force is exerted on the charge carriers because their speed in the rod is zero, which means $\vec{v} \times \vec{B} = \vec{0}$.



field were stationary and the rod were moving to the right (recall Figure 29.1). Any motion of the rod relative to the magnetic field produces the same potential difference across the rod's ends. By the same token, a current is induced in the stationary rectangular loop of **Figure 29.11** by the moving magnetic field source, just as a current is induced in the rectangular loop moving through a constant magnetic field in Figure 29.2. In other words, whenever relative motion occurs between charge carriers and the source of a magnetic field, the resulting rearrangement of charge and potential difference is independent of the choice of inertial reference frame, as are any resulting currents.

What force causes the charge carriers to flow when the loop is stationary? We have encountered two types of forces that are exerted on charged particles: Electric and magnetic. Magnetic forces are exerted only on moving charged particles, so these forces cannot be responsible for the induced current in Figure 29.11. Only electric forces are exerted on stationary charged particles, so we must conclude that the force that causes charge to separate in Figure 29.10 is an electric force

Figure 29.11 A current is induced in a stationary conducting loop as a magnetic field source moves past it when not all of the loop is in the field, even though no magnetic force is exerted on the charge carriers.



CONCEPTS

(because it behaves like one). Because electric forces are caused by electric fields, we are led to conclude:

A changing magnetic field is accompanied by an electric field.

29.4 In Figure 29.1, charge accumulates at the ends of the moving rod until the amount at each end reaches an equilibrium value. Mechanical equilibrium is established when the magnetic force due to the motion of the rod counterbalances the electric force due to the charge separation. In Figure 29.10, what two forces determine the equilibrium state of the charge separation in the stationary rod?

We see that the nature of the force causing electromagnetic induction depends on the choice of reference frame. Forces that appear to be magnetic in one inertial reference frame are electric in another inertial reference frame. Indeed, the fact that our choice of reference frame affects whether a force is magnetic or electric is further evidence that electric and magnetic fields arise from fundamentally the same interaction (see Section 27.4). Because viewing the situation from a different inertial reference frame cannot alter the underlying interaction, magnetic and electric forces must be two manifestations of the same interaction.

29.5 As viewed from above, what is the direction of the induced current in situations 1 and 3 in Figure 29.8?

What do the field lines of the electric field that induces the current in Figure 29.11 look like? Electric field lines show the direction of the electric force exerted on a positively charged test particle. Because a counterclockwise current is induced in the loop (see position *d* in Example 29.1), we know that positive charge carriers in the loop experience a force directed counterclockwise all along the loop. This leads us to conclude that the field lines must form loops that close on themselves. The electric field lines that accompany a changing magnetic field are therefore very different from the electric field lines we have encountered thus far, which originate and terminate on charged objects. Returning to the loop in Figure 29.11, we see that there is no charge separation anywhere in the loop, so there is no particular place where a field line begins or ends.

CONCEPTS

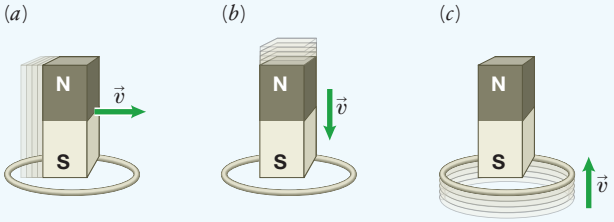
29.4 Lenz's law

What determines the direction of the current induced by a changing magnetic flux? So far, the only way we have to determine the direction of the induced current is to work out the direction of the magnetic forces by considering the situation from a frame of reference in which the magnetic field is stationary and the charge carriers are moving. As the following example illustrates, however, this procedure is cumbersome at best.

EXAMPLE 29.3 What is the current direction?

For each of the three situations shown in **Figure 29.12**, is the induced current in the loop clockwise or counterclockwise as viewed from above?

Figure 29.12 Examples 29.3 and 29.4.



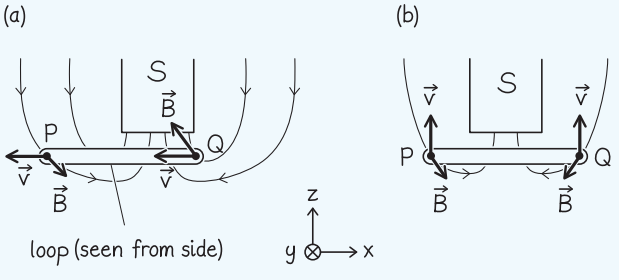
1 GETTING STARTED To determine the direction of the induced current, I need to determine the direction of the electric or magnetic force exerted on the charge carriers in the loop.

2 DEVISE PLAN I can use the right-hand force rule to determine the direction of the magnetic forces (assuming positive charge carriers), and the only way I know to do this is to consider each situation from a reference frame in which the magnetic field is stationary and the loop is in motion. For (*a*) I choose a reference frame moving along with the magnet; in this reference frame the loop moves to the left (**Figure 29.13*a***). Situations (*b*) and (*c*) are equivalent, so the direction of the induced current must be the same in both. I therefore need to consider only (*c*), where the magnetic field is stationary and the loop is in motion (Figure 29.13*b*). I can then use the right-hand force rule to determine the direction of the magnetic force at a couple of locations on the loop. Once I know these force directions, I can determine the direction of the induced current.

3 EXECUTE PLAN I'll consider two locations on the loop: location P on the left and Q on the right. At each of these locations I draw arrows representing the velocity $\vec{v}$ of the (positive) charge carriers and the magnetic field $\vec{B}$. (All of the vectors I draw are in the vertical *xz* plane.)

(*a*) At location P on the loop in Figure 29.13*a*, the magnetic field points down and to the right, and the charge carriers in the loop are moving to the left at velocity $\vec{v}$. If I curl the fingers of my right hand from $\vec{v}$ to $\vec{B}$ at P, my thumb points out of the page, which means the magnetic force exerted on the charge carriers at P points out of the page. Thus at P the charge carriers are moving counterclockwise as viewed from above. At Q the magnetic field is

Figure 29.13



up and to the left, and the charge carriers here are also moving to the left at velocity $\vec{v}$. When I curl the fingers of my right hand from $\vec{v}$ to $\vec{B}$, my thumb points into the page. This means the magnetic force at Q points into the page, moving the charge carriers in the loop counterclockwise as viewed from above. Thus the induced current is counterclockwise as seen from above. ✔

(*b*, *c*) At P in Figure 29.13*b*, the loop and its charge carriers are moving up at velocity $\vec{v}$, and the magnetic field has a vertical component and a component to the right in the plane of the page. Curling the fingers of my right hand from $\vec{v}$ to $\vec{B}$, I discover that the magnetic force exerted on the positive charge carriers points into the page. At Q it points out of the page, so the magnetic forces exerted on the charge carriers at P and Q induce a current that is clockwise as seen from above. ✔

4 EVALUATE RESULT Comparing Figures 29.8 and 29.12, I note that Figure 29.12*a* is similar to situation 1 in Figure 29.8, and Figure 29.12*b* and *c* are similar to situations 3 and 4. In Example 29.2, I concluded that currents are induced in situations 1, 3, and 4, so in this example, too, I should expect currents in all three situations.

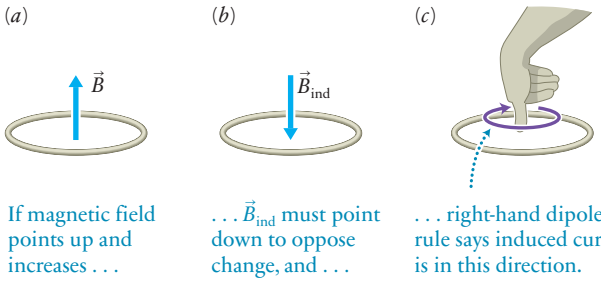
A simpler approach to determining the direction of an induced current follows from experimental observation:

> **The direction of an induced current through a conducting loop is always such that the magnetic flux produced by the induced current opposes the change in magnetic flux through the loop.**

This principle is called **Lenz's law,** and the magnetic field resulting from the induced current is called the **induced magnetic field.**

Consider, for example, a conducting loop in a magnetic field that points up and is increasing in magnitude (**Figure 29.14*a***). In order for the induced magnetic field $\vec{B}_{\text{ind}}$ to oppose the change in magnetic field, $\vec{B}_{\text{ind}}$ must point down, as shown in Figure 29.14*b*. What current direction produces a downward induced magnetic field? According to the right-hand dipole rule, the current must be clockwise as seen from above (Figure 29.14*c*). Indeed, this is what we found for a loop moving into a magnetic field in Example 29.1: As the magnetic flux through the loop increases, a clockwise current is induced.

Figure 29.14 Applying Lenz's law.



EXAMPLE 29.4 Clockwise or counterclockwise?

Consider again the three situations of Figure 29.12. Use Lenz's law to determine the direction of the induced current in each loop.

1 GETTING STARTED Using Lenz's law involves determining which direction for the induced current produces an induced magnetic field opposing the change in magnetic flux through the loop. The magnetic field that passes through the loop points primarily up (toward the south pole of the magnet) and is strongest along the long axis of the magnet.

2 DEVISE PLAN According to Lenz's law, the induced current must be directed in such a way as to create an induced magnetic field $\vec{B}_{\text{ind}}$ that counteracts the change in magnetic flux. If the magnetic flux that causes the induced current decreases, $\vec{B}_{\text{ind}}$ must be in the same direction as the magnetic field of the magnet so that $\vec{B}_{\text{ind}}$ increases the magnetic flux through the loop. If the magnetic flux that causes the induced current increases, $\vec{B}_{\text{ind}}$ must be in the direction opposite the direction of the magnet's magnetic field. Once I know the direction of $\vec{B}_{\text{ind}}$, I can use the right-hand dipole rule to determine the direction of the current.

3 EXECUTE PLAN (*a*) As the magnet moves to the right, the magnitude of the magnetic field at the loop, and therefore the magnetic flux through the loop, decrease. Therefore $\vec{B}_{\text{ind}}$ is upward, in the direction of the magnet's magnetic field. To use the right-hand dipole rule, I recall from Section 28.3 that the direction of a magnetic field near the poles of a bar magnet is given by the magnetic dipole moment vector, which points from south pole to north pole. Thus I point my right thumb in this direction in Figure 29.12a and see that my fingers curl counterclockwise as viewed from above, telling me that this is the direction of the induced current. ✔

(*b*) Moving the magnet toward the loop causes the magnetic flux through the loop to increase. The induced magnetic field $\vec{B}_{\text{ind}}$ associated with the induced current must therefore point downward, opposite the direction of the field created by the magnet, to oppose the increase in magnetic flux through the loop. The induced current is therefore clockwise as viewed from above. ✔

(*c*) Moving the loop toward the magnet results in an increase in magnetic flux through the loop, as in part *b*. The induced current therefore must be clockwise as viewed from above to produce a downward-pointing induced magnetic field $\vec{B}_{\text{ind}}$ that opposes the increase in flux. ✔

4 EVALUATE RESULT My answers are the same as in Example 29.3, which gives me confidence that they are correct.

29.6 When current is induced in a conducting loop by the motion of a nearby magnet, the induced magnetic field $\vec{B}_{\text{ind}}$ exerts a force on the magnet. (*a*) In Figure 29.12*b*, what is the direction of the force exerted on the magnet by $\vec{B}_{\text{ind}}$? (*b*) Suppose Lenz's law stated that the induced current *adds to* the change that produced it instead of opposing it. What would be the direction of the force exerted by $\vec{B}_{\text{ind}}$ on the magnet?

CONCEPTS

Checkpoint 29.6 gives us a clue to the reason for Lenz's law. In Figure 29.12*b*, let's consider the system made up of magnet and loop. If the induced current further increased the magnetic flux through the loop, the induced current would exert an attractive force on the magnet. This attractive force would pull the magnet closer to the loop, increasing the magnetic flux further and inducing even more current. In the process, the kinetic energy of the magnet and the energy associated with the current in the loop would increase, increasing the energy of the magnet-loop system without any work being done on it, violating the law of conservation of energy.

Conservation of energy therefore *requires* that an induced current oppose the change that created it. In Figure 29.12*b*, for example, the force exerted by the induced magnetic field $\vec{B}_{\text{ind}}$ resists the magnet moving closer to the loop. An alternative statement of Lenz's law therefore is:

An induced current is always in such a direction as to oppose the motion or change that caused it.

Because of this opposing force, an agent moving the magnet at constant speed toward the loop must do work on the magnet. Where does this energy go? As work is done on the system, the current through the loop—and therefore the induced magnetic field—increases. Just as we can associate electric potential energy with an electric field, we can associate **magnetic potential energy** with a magnetic field. In the absence of dissipation, the work done on the magnet-loop system therefore increases the magnetic potential energy of the system. The energy diagram in **Figure 29.15** illustrates the conversion of mechanical work to magnetic potential energy.

Figure 29.15 Energy diagram for a magnet-loop system when the magnet moves toward the conducting loop, inducing a current through the loop. Work must be done on the system to push the magnet at constant speed against the opposing induced magnetic field of the loop. This work causes the potential energy of the magnet-loop system to increase.

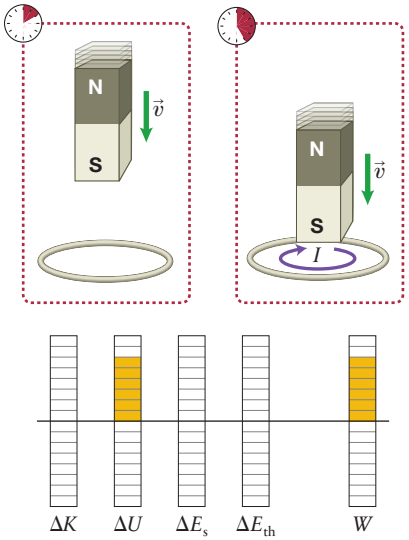
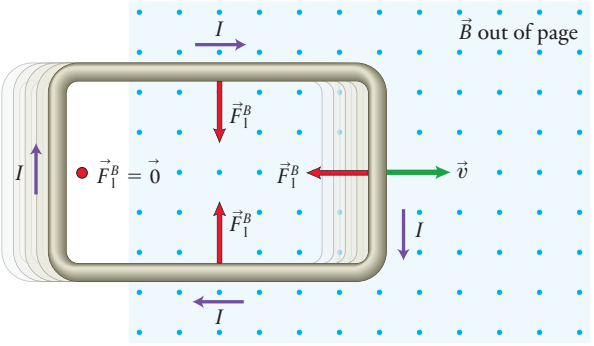


Figure 29.16 Direction of magnetic forces exerted by a magnetic field on each side of a rectangular conducting loop because of the current induced in the loop. As the loop moves into the field, the vector sum of the magnetic forces exerted on the loop resists its motion into the field.



Work must also be done to move a conducting loop into a magnetic field. When current is induced in the loop by its motion in the magnetic field, the magnetic field exerts a force on the current.

Figure 29.16 shows the directions of these magnetic forces exerted on each side of a rectangular conducting loop as it enters a magnetic field from the left. The vector sum of the forces points to the left. In order to move the loop through the field at constant speed, you must exert on the loop a constant force to the right that is equal in magnitude to the magnetic force (**Figure 29.17**). In doing so, you do work on the loop.

CONCEPTS

29.7 (*a*) After the left edge of the loop in Figure 29.17 enters the magnetic field, is the work required to continue pulling the loop through the field at constant speed v positive, negative, or zero? (*b*) As the right edge emerges from the magnetic field, is the work required positive, negative, or zero?

Figure 29.17 Work must be done on a conducting loop to move it into a magnetic field because magnetic forces exerted on the loop resist its motion when there is an induced current in the loop (shown in more detail in Figure 29.16).

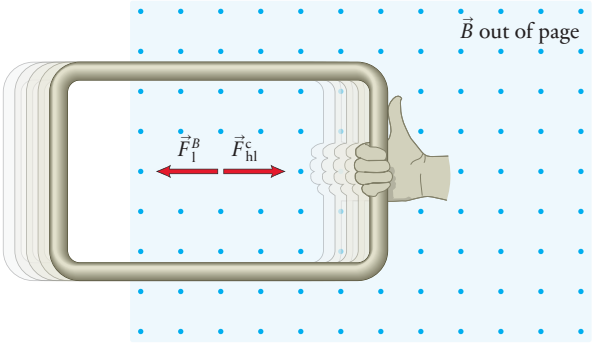
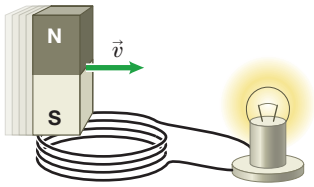


Figure 29.18 An induced current can be used to light a lamp. The work done on the magnet is converted to light and thermal energy in the filament of the lamp.



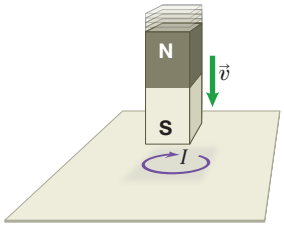
As you move a conductor relative to a magnetic field (or vice versa), the work you do transfers energy to the conductor, setting the charge carriers in motion. This energy associated with the induced current can then be converted to other forms of energy by some device placed in the current path, such as a lamp, toaster, or electric motor. In **Figure 29.18**, for example, a lamp is inserted in the current path and the work done on the system is converted to light energy and thermal energy in the lamp. Doing work by moving a loop through a magnetic field and thereby setting charge carriers in motion is the basic idea behind electric generators.

29.8 Which requires doing more work: moving a magnet toward a closed conducting loop or moving it toward a rod? Both motions are at constant speed.

It is not necessary to have a conducting object in the shape of a loop. If a bar magnet is moved toward an extended conducting object, such as a conducting sheet (**Figure 29.19**), circular currents are induced throughout the surface of the object. These currents are called **eddy currents** because they form eddy-like loops in the object's surface. (An *eddy* is the circular, whirlpool-like movement of water.)

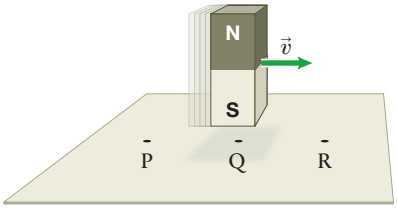
Because eddy currents dissipate energy, they can be used to convert kinetic energy (for example, the kinetic energy associated with the motion of a conductor relative to a magnet) to thermal energy. Dissipation of energy by eddy currents is used in braking systems, especially trains and roller coasters. In contrast to conventional friction brakes, eddy current brakes involve no physical contact and so there is no wear and tear and there are no brake pads to be replaced.

Figure 29.19 The circular current loops induced in a conducting sheet by the motion of a nearby magnet are called eddy currents.



29.9 In **Figure 29.20**, a bar magnet moves parallel to a metal plate. (*a*) At the instant shown, does the magnitude of the magnetic flux increase, decrease, or stay the same through a small region around points P, Q, and R? (*b*) Are the eddy currents induced around these points clockwise, counterclockwise, or zero?

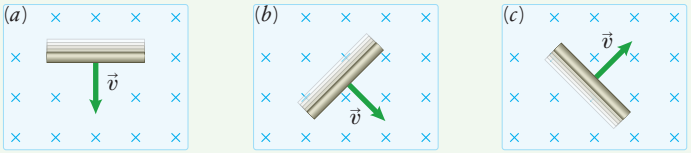
Figure 29.20 A bar magnet moves parallel to the surface of a conducting sheet, causing the local magnetic flux to change across the sheet surface and inducing eddy currents where the magnetic flux is changing (Checkpoint 29.9).



Self-quiz

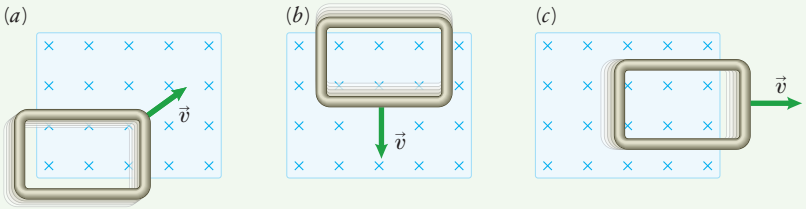
1. A conducting rod moves through a magnetic field as shown in **Figure 29.21**. Which end of the bar, if any, becomes positively charged?

Figure 29.21



2. A conducting loop moves through a magnetic field as shown in **Figures 29.22a–c**. Which way does the current run in the loop at the instant shown in each figure?

Figure 29.22



3. A conducting loop moves through a magnetic field at constant velocity as shown in **Figure 29.23**. For each case *a*–*e*, must the work done on the loop be positive, negative, or zero to keep the loop moving?
4. Using Faraday's law, determine whether charge carriers flow in the loop for each situation shown in **Figure 29.24**.

Figure 29.23

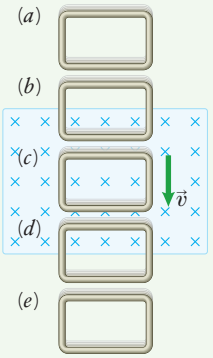
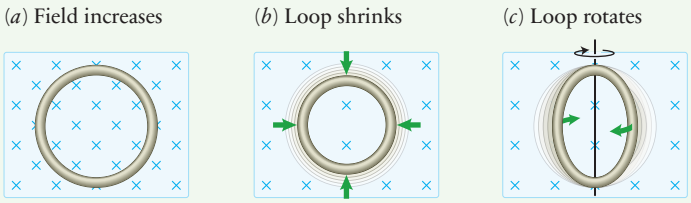


Figure 29.24



5. Using Lenz's law, determine the direction of the induced current, if any, at the instants shown in Figure 29.24.

Answers

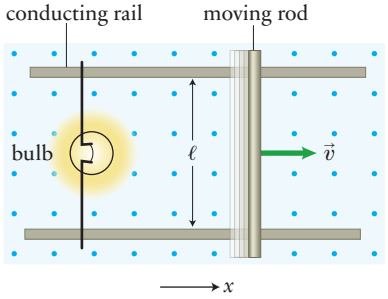
1. The direction of the magnetic force on the positive charge carriers in the rod is given by the right-hand force rule: If you point the fingers of your right hand in the direction of $\vec{v}$ and then curl them toward $\vec{B}$ (into the page), your thumb points in the direction of the force. Applying this rule in the three situations shows that positive charge accumulates on the right end of the bar in (*a*) and (*b*) and on the left end of the bar in (*c*).
2. In (*a*) and (*b*), the magnitude of the magnetic flux through the loop increases into the page, so the induced magnetic field opposing this change must point out of the page. According to the right-hand dipole rule, an induced current must be counterclockwise to produce a magnetic field that points out of the page. In (*c*), the magnitude of the magnetic flux through the loop decreases into the page, so the induced current must be clockwise.
3. In (*a*), (*c*), and (*e*), the magnetic flux through the loop does not change and so no current is induced in the loop. The vector sum of the magnetic forces exerted on the loop is therefore zero, so no work is required to keep the loop moving at constant speed. In (*b*) and (*d*), the magnetic flux through the loop changes and a current is induced through the loop. According to Lenz's law the induced current is in such a direction as to oppose the motion or change that created it, so a positive amount of work is needed to keep the loop moving at constant speed.
4. Because the magnetic flux through the loop is changing in all three cases, charge carriers flow in all three.
5. (*a*) The magnetic field is directed into the page, and the magnitude of the magnetic flux is increasing into the page. The induced magnetic field opposing this change must therefore point out of the page, so according to the right-hand dipole rule, the induced current is counterclockwise. (*b*) Because the area of the loop decreases, the magnetic field is directed into the page and the magnitude of the magnetic flux decreases. The induced current is therefore clockwise. (*c*) At the instant shown in the figure, the magnitude of the magnetic flux through the loop decreases and so the induced current must be clockwise.

29.5 Induced emf

We first encountered the concept of emf in Chapter 26 in the context of batteries. Chemical reactions in a battery produce and maintain a separation of charge across the battery terminals; this separation results in a potential difference from one terminal to the other. If a light bulb is connected to the terminals, this potential difference drives a current through the light bulb. In Section 26.4, we defined the emf of a charge-separating device as the work done per unit charge by nonelectrostatic interactions in separating charge within the device. Because the charge separation across the ends of a rod caused by electromagnetic induction (Figure 29.10) is nonelectrostatic, we can associate an emf with this separation of charge. Emfs produced by electromagnetic induction are therefore called **induced emfs.**

Consider the setup illustrated in **Figure 29.25**. A conducting rod of length ℓ rests on two conducting rails connected to a light bulb. The rod is pulled to the right and moves at a constant speed v along the x axis. The motion of the rod changes the area of the magnetic field enclosed by the loop formed by the rod, rails, and wires attaching the bulb to the rails. Because of this area change, a current is induced through the loop. This current causes the light bulb to glow.

Figure 29.25 The emf that develops in a moving rod can be used to drive a current when that rod is part of a closed conducting loop. The rod is connected to the rest of the loop by sliding electrical contacts.



As we saw in Section 29.4, the magnetic field exerts a leftward force on the current-carrying rod. To move the rod at a constant speed v to the right, therefore, the agent pulling the rod must exert a rightward force of equal magnitude on it. In doing so, the external agent must do work on the rod, increasing the energy of the charge carriers in the loop. This energy is converted to light and thermal energy in the light bulb. We can obtain an expression for the induced emf in the loop by calculating the work done on the rod.

According to Eq. 27.4, the magnitude of the magnetic force exerted on a rod of length ℓ carrying a current I is $|I|\ell B$ when the rod is perpendicular to the magnetic field, so the contact force $\vec{F}^{\text{c}}_{\text{ar}}$ exerted by the agent on the rod required to pull it at constant speed must be of the same magnitude: $F^{\text{c}}_{\text{ar}} = |I|\ell B$. Because the work done on the rod is positive, it is equal to the product of the magnitude of the contact force exerted on it and the magnitude of the force displacement: $W_{\text{r}} \equiv F^{\text{c}}_{\text{ar}x}\Delta x_F = F^{\text{c}}_{\text{ar}}|\Delta x_F|$. In a time interval Δt, the rod is displaced a distance $\Delta x = v_x \Delta t$, so we have

$$W_{\text{r}} = F^{\text{c}}_{\text{ar}}|\Delta x_F| = B|I|\ell v \Delta t \tag{29.1}$$

or, using Eq. 27.1, $I \equiv q/\Delta t$, to simplify,

$$W_{\text{r}} = B\ell v|q|, \tag{29.2}$$

where q is the charge that passes through a given section of the rod in a time interval Δt. This result tells us how much energy is transferred to the current loop as the rod is pulled. Note that the force exerted by the magnetic field on the charge carriers in the rod transfers the mechanical work done on the rod to the charge carriers in the loop—the magnetic force itself does no work. (Because the direction of the magnetic force is always perpendicular to the motion of the charge carriers, the force cannot do any work on the carriers.)

If it bothers you that the external agent is doing the work and not the magnetic force that sets the charge carriers in motion, consider an analogy: Suppose you push yourself away from a wall starting from rest. The wall provides the force that accelerates you away from it. This force does *no work*, however, because the point of application does not move. In fact, no work at all is done on you. Internal (source) energy is converted to kinetic energy; the force exerted by the wall on you converts the internal energy to kinetic energy.

We can now calculate the magnitude of the induced emf, which is defined as the nonelectrostatic work per unit charge done on a charged particle (Eq. 26.7):

$$|\mathscr{E}_{\text{ind}}| \equiv \frac{W_{\text{r}}}{|q|} = B\ell v. \tag{29.3}$$

As shown in Figure 29.25, the induced emf in the closed conducting loop drives a current, just as a battery of the same emf would. This current has the same effect on the loop as a battery-driven current would have; in this case, it lights a light bulb. The magnitude of the resulting current—the **induced current**—depends on how much resistance the loop puts up to the motion of charge carriers. Experimentally we determine that the induced current through many conducting materials is proportional to the induced emf:

$$I_{\text{ind}} = \frac{\mathscr{E}_{\text{ind}}}{R}, \tag{29.4}$$

where the proportionality constant R is a measure of the *resistance* of the loop. The smaller the value of R, the greater the induced current for a given induced emf. We'll learn more about resistance in Chapter 31.

Equation 29.3 gives the magnitude of the induced emf in terms of v, B, and the length ℓ of the rod. In Section 29.2 we argued that the induced current—that is, the current that results from the induced emf—is due to a change in the magnetic flux enclosed by the loop. Let us therefore obtain an expression for the induced emf in terms of the magnetic flux Φ_B through the loop. In Section 27.6 we defined the magnetic flux through a loop by the surface integral (Eq. 27.10)

$$\Phi_B \equiv \int \vec{B} \cdot d\vec{A}, \tag{29.5}$$

where $d\vec{A}$ is the area vector of an infinitesimally small segment of a surface bounded by the loop.

Because the surface integral in Eq. 29.5 is independent of the choice of surface bounded by the loop (Eq. 27.15), we integrate over a flat surface bounded by the loop in Figure 29.25. Because the magnetic field is uniform, we can pull it out of the integral and so in Figure 29.25, if we let the vector $d\vec{A}$ point out of the page in the same direction as $\vec{B}$, $\Phi_B = BA$. When the moving side of the loop is displaced by Δx, the area enclosed by the expanding loop changes by $\Delta A = \ell \Delta x$. If this change occurs over a time interval Δt, the rate of change in the magnetic flux enclosed by the loop is thus

$$\frac{\Delta \Phi_B}{\Delta t} = \frac{B \Delta A}{\Delta t} = \frac{B\ell \Delta x}{\Delta t} = B\ell \frac{\Delta x}{\Delta t}, \tag{29.6}$$

so Eq. 29.3 can be written as

$$|\mathscr{E}_{\text{ind}}| = B\ell v = B\ell \left| \frac{\Delta x}{\Delta t} \right| = \left| \frac{\Delta \Phi_B}{\Delta t} \right|. \tag{29.7}$$

As we let the time interval Δt approach zero, this yields

$$\mathscr{E}_{\text{ind}} = -\frac{d\Phi_B}{dt}, \tag{29.8}$$

QUANTITATIVE TOOLS

where we have added a negative sign to indicate that the direction of the induced emf is such that it drives a current that counteracts the change in magnetic flux, in accordance with Lenz's law. Equation 29.8, which relates a changing magnetic flux to an induced emf, is a quantitative statement of **Faraday's law.**

Note that Eq. 29.8 correctly gives a zero emf when applied to the situation shown in Figure 29.4*c*, in which the entire loop (not just one side) is moving and is in the magnetic field. We derived Eq. 29.8 for the special case of one side of a rectangular conducting loop moving in a magnetic field, but it is generally valid: It gives the value of the emf no matter what causes the magnetic flux through a conducting loop to change.

29.10 Sketch how the induced emf in the loop in Figure 29.4 varies as the loop moves through the five positions.

Equation 29.8 allows us to point out an important difference between potential difference and emf, both of which are related to the work per unit charge done on charge carriers—the former by electrostatic interactions, the latter by nonelectrostatic interactions. As we established in Chapter 25, potential difference depends only on starting and ending locations. It does not depend on the path followed to get from the starting location to the ending location. In other words, potential difference is path-independent. The work per unit charge done on charged particles by *nonelectrostatic* interactions, such as the induced emf in Figure 29.25, does depend on the path taken. Consider, for example, a rectangular conducting loop containing a light bulb and placed in a changing magnetic field (**Figure 29.26*a***). The light bulb glows because the changing magnetic flux through the area enclosed by the loop causes an induced emf throughout the loop. If the area enclosed by the loop is smaller (Figure 29.26*b*), then the light bulb is less bright. The emf induced in the loop is now smaller because the smaller area causes a smaller rate of change in the magnetic flux. Had we established a potential difference across the light bulb by connecting it via wires to, say, the terminals of a battery, how much area the wires enclose would not affect the brightness of the bulb (**Figure 29.27**). We shall discuss this difference in path-dependence between potential difference and emf in more detail in the next section.

We can use an induced emf to establish a potential difference between two locations. Consider, for example, the conducting rod of Figure 29.1. As we saw in Section 29.1, the magnetic force exerted on the negative charge carriers in the moving rod drives them upward. A surplus of negative charge accumulates at the top of the rod, leaving a surplus of positive charge at the bottom. This separation of charge produces an electric field inside the rod. The negative charge carriers flow until the magnitude of the downward electric force exerted on one of them due to the electric field inside the rod is equal to the magnitude of the upward magnetic force (**Figure 29.28** on the next page):

$$F_q^E = F_q^B. \tag{29.9}$$

The magnitude of the magnetic force exerted on a charged particle of charge q moving perpendicular to a magnetic field is given by Eq. 27.18, $F_q^B \equiv |q|vB$, and the electric force is given by Eq. 23.6, $F_q^E = |q|E$. With this information, we can rewrite Eq. 29.9 as

$$|q|E = |q|vB. \tag{29.10}$$

Thus E, the magnitude of the electric field inside the rod, is

$$E = vB. \tag{29.11}$$

Figure 29.26 Path-dependence of an induced emf.

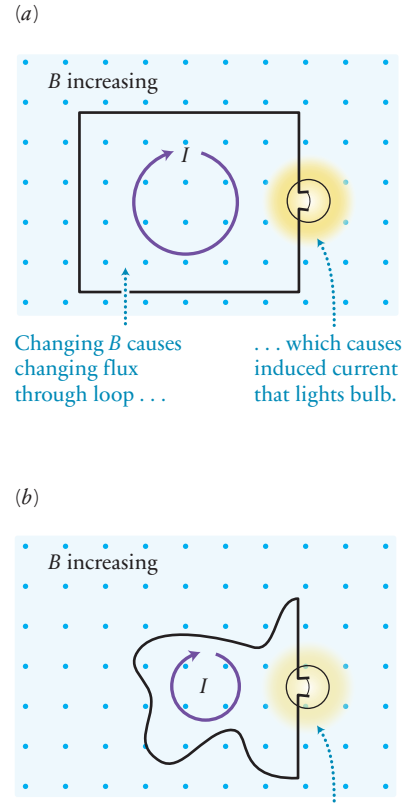
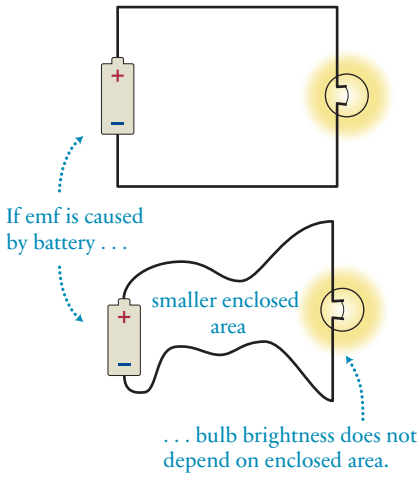
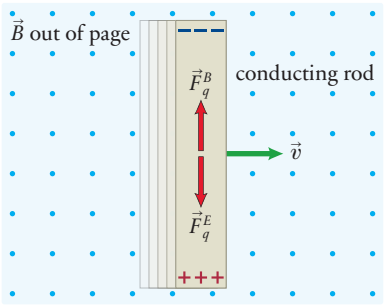


Figure 29.27 The emf produced by a battery does not depend on enclosed area.



QUANTITATIVE TOOLS

Figure 29.28 When a conducting rod moves in a magnetic field, the magnetic force exerted on the negative charge carriers in the rod causes a charge separation. This separation continues until the force exerted by the electric field resulting from the charge separation exactly counters the magnetic force exerted on the charge carriers.



This electric field causes a potential difference between the top and the bottom of the rod. For a rod of length ℓ, the potential difference is (see Eq. 25.25)

$$V_{\text{tb}} = -\int_{\text{top}}^{\text{bottom}} \vec{E}\cdot d\vec{\ell} = \int_{\text{top}}^{\text{bottom}} E\,d\ell = E\int_{\text{top}}^{\text{bottom}} d\ell = E\ell, \qquad (29.12)$$

where $\vec{E}\cdot d\vec{\ell} = -E\,d\ell$ because the electric field points from the bottom to the top of the rod whereas $d\vec{\ell}$ points in the direction of integration. In addition we have pulled E out of the integral because it is constant throughout the rod. Substituting the magnitude of the electric field from Eq. 29.11, we obtain

$$V_{\text{tb}} = vB\ell. \qquad (29.13)$$

Thus we see that the electric field in the rod establishes across the ends of the rod a potential difference that is equal in magnitude to the induced emf (Eq. 29.3).

EXAMPLE 29.5 Airplane wing "battery"

A Boeing 747 with a wingspan of 60 m flies at a cruising speed of 850 km/h. What is the magnitude of the maximum potential difference induced between the two wingtips by Earth's magnetic field (the magnitude of which is roughly 0.50×10^{-4} T)?

❶ **GETTING STARTED** A Boeing 747 is made of metal and so, as the airplane flies through the magnetic field of Earth, the electrons in the metal are able to move when a magnetic force is exerted on them. This motion of the electrons causes a charge separation in the wings, and thus a potential difference is induced. To solve this problem, I must make several simplifying assumptions. I take the wings to be a 60-m conducting rod and assume the plane is flying horizontally through a vertical magnetic field. As a result, the problem reduces to a rod moving through a perpendicular magnetic field.

❷ **DEVISE PLAN** To calculate the induced potential difference between the ends of the "rod" formed by the wings, I can use Eq. 29.13.

❸ **EXECUTE PLAN**

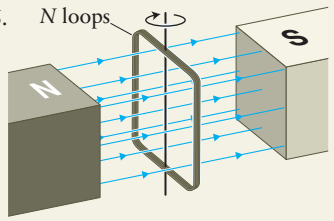
$$V_{\text{wings}} = (8.50 \times 10^5 \text{ mh})\left(\frac{1\text{ h}}{3600\text{ s}}\right)(0.50 \times 10^{-4}\text{ T})(60\text{ m})$$
$$= 0.71\text{ T}\cdot\text{m}^2/\text{s} = 0.71\text{ V}. \checkmark$$

❹ **EVALUATE RESULT** My calculation yields a potential difference on the order of the one from an AA battery, but that should not cause any problems in the aircraft. Also, the magnetic field is unlikely to be perpendicular to the velocity of the plane, as I assumed, and therefore the actual value of the potential difference is smaller than what I obtained.

EXAMPLE 29.6 Generator

In an electric generator a solenoid that contains N windings each of area A is rotated at constant rotational speed ω in a uniform magnetic field of magnitude B (**Figure 29.29**). What is the emf induced in the solenoid?

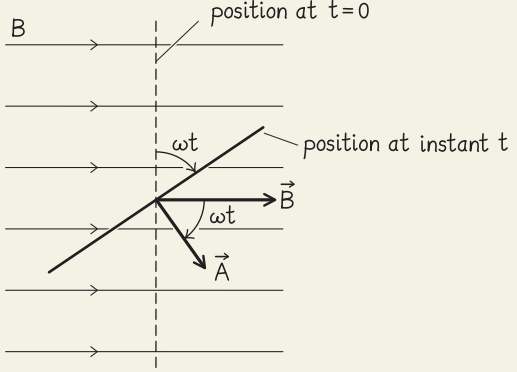
Figure 29.29 Example 29.6.



❶ **GETTING STARTED** As the solenoid rotates, the magnetic flux through it changes, and this changing flux causes an emf in the solenoid.

❷ **DEVISE PLAN** Equation 29.8 tells me that the emf induced in the solenoid equals the time rate of change of the magnetic flux through it. Therefore I must first determine how the magnetic flux varies as a function of time and then differentiate whatever expression I get to obtain the emf. I can determine the magnetic flux through the solenoid by multiplying the magnetic flux through a single winding, $\Phi_B \equiv \vec{B}\cdot\vec{A}$ (Eq. 27.9), by the number of windings N. To determine the magnetic flux through a single winding, I sketch a top view of a winding in the magnetic field, indicating the directions of the magnetic field $\vec{B}$ and the area vector $\vec{A}$ (**Figure 29.30**). I let the plane of the winding be perpendicular to the direction of the magnetic field at $t = 0$.

Figure 29.30



QUANTITATIVE TOOLS

3 EXECUTE PLAN As the solenoid rotates, the scalar product $\vec{B} \cdot \vec{A}$ changes with time. At instant t shown in my sketch, the angle between $\vec{A}$ and $\vec{B}$ is ωt, and so the magnetic flux through a single winding is $\Phi_B = \vec{B} \cdot \vec{A} = BA \cos \omega t$. Through the N windings of the solenoid the magnetic flux is $\Phi_B = NBA \cos \omega t$. Substituting this value into Eq. 29.8, I get

$$\mathcal{E}_{\text{ind}} = -\frac{d}{dt}(NBA \cos \omega t) = \omega NBA \sin \omega t. \checkmark$$

4 EVALUATE RESULT My result shows that the emf oscillates sinusoidally. It is zero when $\omega t = n\pi$ $(n = 0, 1, 2, \ldots)$ and maximum when $\omega t = n\pi + \frac{\pi}{2}$ $(n = 0, 1, 2, \ldots)$. That result makes sense because the rate of change of the magnetic flux through the solenoid is zero when the area vector of the windings is parallel to the magnetic field $(\omega t = n\pi)$ and maximum when the area vector is perpendicular to the magnetic field $(\omega t = n\pi + \frac{\pi}{2})$.

29.11 The expression I derived in Example 29.6 indicates that the emf becomes negative after the solenoid has rotated 180° and remains negative through the next 180° of rotation. However, the solenoid orientation looks the same when the solenoid has rotated 180° as when it started. Why does the emf have a different sign for half of the rotation?

29.6 Electric field accompanying a changing magnetic field

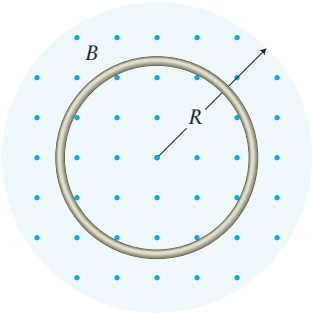
We saw in Section 29.3 that when the magnetic flux through a conducting loop changes, an electric force is exerted on the initially stationary charge carriers in the loop. This electric force is caused by the electric field that accompanies the changing magnetic field. Let's explore the properties of this electric field in more detail.

Consider the case of a conducting circular loop in a uniform magnetic field that has a circular cross section (**Figure 29.31a**). Suppose the magnitude of the magnetic field increases steadily over time. Because the conducting loop encloses an increasing magnetic flux directed out of the page, a clockwise current is induced in the loop. What do the electric field lines that are responsible for this current look like?

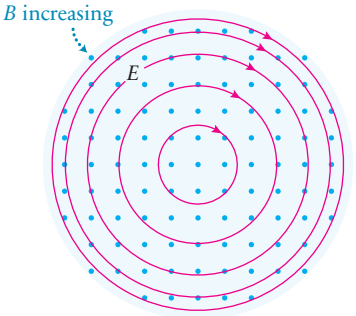
We found in Section 29.3 that the electric field lines that accompany a changing magnetic field form loops because there are no isolated charge carriers on which the electric field lines can begin or end. These loops must be circular and centered on the axis of the magnetic field because the electric field that accompanies the changing magnetic field must have the same cylindrical symmetry as the magnetic field (Figure 29.31*b*). Lenz's law tells us that the induced current is clockwise, and so the electric field must also be pointing clockwise. When the magnetic field does not exhibit cylindrical symmetry, the electric field lines are not circular, and in general it is difficult to determine the shape of the electric field lines.

Figure 29.31 Electric field that accompanies an increasing cylindrical magnetic field.

(*a*) Conducting ring in cylindrical uniform magnetic field



(*b*) Electric field that accompanies increasing magnetic field



29.12 What do the electric field lines look like when the magnitude of the magnetic field in Figure 29.31*b* (*a*) is held constant and (*b*) decreases steadily?

To determine the magnitude of the electric field that accompanies the changing magnetic field, let's think about its effect on charge carriers in the conducting loop. A current is induced in the loop because the electric field does work on the charge carriers in the loop. We can calculate the work done by the electric field using Eq. 25.1, which gives the work done by an electric field $\vec{E}$ on a particle carrying a charge q in moving it from point A to point B:

$$W_q(\text{A} \to \text{B}) = \int_{\text{A}}^{\text{B}} \vec{F}_q^E \cdot d\vec{\ell} = q \int_{\text{A}}^{\text{B}} \vec{E} \cdot d\vec{\ell}. \qquad (29.14)$$

QUANTITATIVE TOOLS

(Note that the electric field accompanying the changing magnetic field is *not* an electrostatic field, and thus the quantity calculated in Eq. 29.14 is *not* electrostatic work.)

The work done on a particle carrying a charge q as it travels around the closed path formed by the conducting loop is then

$$W_q \text{ (closed path)} = q \oint \vec{E} \cdot d\vec{\ell}. \tag{29.15}$$

The work done by the electric field per unit charge is thus

$$\frac{W_q}{q} = \oint \vec{E} \cdot d\vec{\ell}. \tag{29.16}$$

This work per unit charge is the induced emf (Eq. 26.7), and so combining Eq. 29.16 with Eq. 29.8 gives us an expression that contains the electric field accompanying the changing magnetic field:

$$\oint \vec{E} \cdot d\vec{\ell} = -\frac{d\Phi_B}{dt}. \tag{29.17}$$

Equation 29.17 is an alternative statement of Faraday's law. Note that the magnetic flux appearing in this equation is a signed quantity. In order to obtain the correct direction of the electric field in Eq. 29.17, the magnetic flux must be taken to be positive when the magnetic field points in the direction of your right-hand thumb as you curl the fingers of your right hand in the direction of the integration path.

Keep in mind that the electric field in Eq. 29.17 is not an electrostatic field. Electrostatic fields originate directly from static charge distributions. Rather than originating from static charged objects, the electric field in Eq. 29.17 appears with a changing magnetic field and the electric field lines close on themselves. As we saw in Chapter 25, the work done by an electrostatic field on a charged particle moving around a closed path is zero, which means the line integral around a closed path on the left in Eq. 29.17 is always zero for an electrostatic field (see Eq. 25.32). However, as Eq. 29.17 shows, in a region where the magnetic field is changing, the line integral of an electric field around a closed path is not necessarily zero. This means that the electric field accompanying the changing magnetic field can do work on charged particles that travel along a closed path and the amount of work depends on the choice of path. We cannot define a potential difference between two points for such an electric field because it would have different values for different paths between those points.

EXAMPLE 29.7 Electric field magnitude

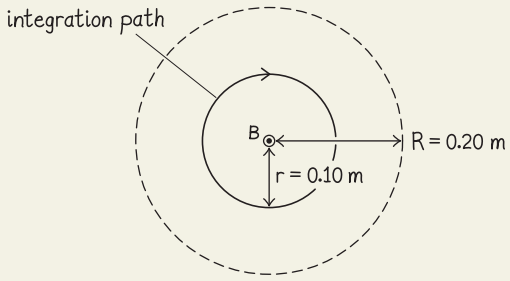
Let the uniform cylindrical magnetic field in Figure 29.31 have a radius $R = 0.20$ m and increase at a steady rate of 0.050 T/s. What is the magnitude of the electric field at a radial distance $r = 0.10$ m from the center of the magnetic field?

❶ **GETTING STARTED** The changing magnetic field is accompanied by an electric field that has circular field lines pointing clockwise (Figure 29.31). Because of the circular symmetry, I know that the magnitude E of the electric field cannot vary along any given electric field line.

❷ **DEVISE PLAN** To solve this problem, I can use Eq. 29.17 to work out the relationship between E and the rate of change of the magnetic field magnitude dB/dt. Because E is constant along any electric field line, it must be constant on *any* circular path centered on the axis of the cylindrical magnetic field. I therefore choose as my path of integration a clockwise circular path of radius $r = 0.10$ m centered on the long central axis of the magnetic field (**Figure 29.32**). Because $\vec{E}$ is always parallel to $d\vec{\ell}$, $\vec{E} \cdot d\vec{\ell} = E\,d\ell$ on the left side of Eq. 29.17, and because the magnitude of the electric field is the same everywhere on the path, I can take E out of the integral. To evaluate the right side, I must first calculate the magnetic flux through the circular integration path. Because the magnetic field is uniform, the magnetic flux is given by Eq. 27.9.

QUANTITATIVE TOOLS

Figure 29.32



❸ **EXECUTE PLAN** The left side of Eq. 29.17 becomes

$$\oint \vec{E} \cdot d\vec{\ell} = 2\pi r E. \tag{1}$$

The magnetic flux through the area enclosed by the integration path is negative (curling the fingers of my right hand in the direction of the integration path makes my thumb point in the direction opposite the direction of $\vec{B}$), and so

$$\Phi_B = -BA = -\pi r^2 B.$$

Thus the right side of Eq. 29.17 becomes

$$-\frac{d\Phi_B}{dt} = +\pi r^2 \frac{dB}{dt}. \tag{2}$$

Substituting Eqs. 1 and 2 into Eq. 29.17 then yields

$$2\pi r E = \pi r^2 \frac{dB}{dt}$$

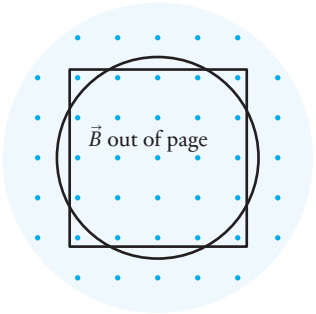
$$E = \tfrac{1}{2} r \frac{dB}{dt} = \frac{(0.10\ \text{m})(0.050\ \text{T/s})}{2}$$

$$= 2.5 \times 10^{-3}\ \text{T} \cdot \text{m/s} = 2.5 \times 10^{-3}\ \text{N/C}. \checkmark \tag{3}$$

❹ **EVALUATE RESULT** Equation 3 shows that the magnitude of the electric field increases with radial distance r from the center of the magnetic field. Given the cylindrical symmetry of the situation, that's exactly what I expect.

Although the electric field lines and the conducting loop pictured in Figure 29.31 are circular, our derivation of Eq. 29.17 does not depend on the shape of the path of integration. This means we can use Eq. 29.17 to determine the line integral of the electric field around any closed path. For example, **Figure 29.33** shows a square path of integration that encloses the same area (and hence the same magnetic flux) as the circular path I used to solve Example 29.7. Because the enclosed magnetic flux is the same, we know that the integral of the electric field around the square path must yield the same result as the integral around the round path, even though the electric field is no longer tangent to the path and no longer of the same magnitude all along the path.

Figure 29.33 Square and circular paths of integration that enclose the same area and therefore enclose the same magnetic flux.





29.13 In Example 29.7 what is the magnitude of the electric field at a distance of 0.30 m from the center of the magnetic field?

29.7 Inductance

An interesting consequence of Faraday's law is that when a current through a conducting loop changes, the change induces an emf *in the loop itself*. To understand why this happens, consider **Figure 29.34*a*** on the next page, in which a battery drives a current I through a solenoid. Because the current creates a magnetic field, a nonzero magnetic flux passes through the loops of the solenoid.

Now imagine the battery is suddenly disconnected, as in Figure 29.34*b*. Without the battery to supply an emf, the current, the magnitude of the magnetic field, and the magnitude of the magnetic flux all start to decrease. Faraday's law tells us, however, that the changing flux induces in the solenoid an emf that opposes the change in flux. Consequently, as the magnetic flux is decreasing, the induced emf causes an induced current that has the same direction as the original current. The magnetic flux associated with this induced current opposes the change in the magnetic flux.

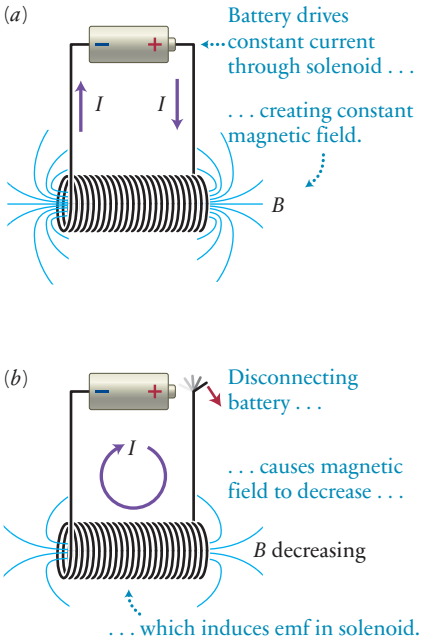


Figure 29.34 Disconnecting a battery from a solenoid induces in the solenoid an emf that opposes the decrease in current through it.

The induced emf is proportional to the rate of change in the magnetic flux through the solenoid (Eq. 29.8). The magnetic flux is proportional to the magnitude B of the magnetic field, which is proportional to the current through the loop (Eq. 28.6, $B = \mu_0 nI$). Thus the faster the current changes, the greater the magnitude of the induced emf:

$$\mathcal{E}_{\text{ind}} \propto \frac{d\Phi_B}{dt} \propto \frac{dB}{dt} \propto \frac{dI}{dt}. \tag{29.18}$$

We can define the **inductance** L of a loop or solenoid as the constant of proportionality between the emf and the rate of change of current:

$$\mathcal{E}_{\text{ind}} = -L\frac{dI}{dt}. \tag{29.19}$$

A large inductance leads to a large induced emf. The derived SI unit of inductance is the **henry:**

$$1\ \text{H} \equiv 1\ \text{V}\cdot\text{s/A} = 1\ \text{kg}\cdot\text{m}^2/\text{C}^2. \tag{29.20}$$

A device that has an appreciable inductance is called an **inductor.** Inductors are used widely in electric circuits to slow down the timing of variations in current.

The inductance describes how much change in magnetic flux is associated with a change in current for a particular loop or solenoid, as we can see by substituting Eq. 29.8 in the left side of Eq. 29.19:

$$\frac{d\Phi_B}{dt} = L\frac{dI}{dt}. \tag{29.21}$$

The inductance of a current-carrying device or circuit depends only on its geometry, but for most real devices, calculating the inductance is not simple. The general procedure for determining the inductance is described in the Procedure box "Calculating inductances" on this page. In a few particularly simple cases, it is possible to derive an algebraic expression for the inductance, as in Example 29.8.

PROCEDURE: Calculating inductances

The inductance of a current-carrying device or current loop is a measure of the emf induced in the device or loop when current is changed. To determine the inductance of a particular device or current loop, follow these four steps.

1. Derive an expression for the magnitude of the magnetic field in the current-carrying device or current loop as a function of the current. Your expression should depend only on the current I and possibly—but not necessarily—the position within the device or current loop.
2. Calculate the magnetic flux Φ_B through the device or current loop. If the expression you derived in step 1 depends on position, you will have to integrate that expression over the volume of the device or circuit. Use symmetry to simplify the integral and divide the device into segments on which B is constant.
3. Substitute the resulting expression you obtained for Φ_B into Eq. 29.21. As you take the derivative with respect to time, keep in mind that only the current varies with respect to time, so you should end up with an expression that contains the derivative dI/dt on both sides of the equal sign.
4. Solve your expression for L after eliminating dI/dt.

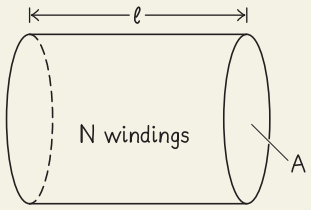
QUANTITATIVE TOOLS

EXAMPLE 29.8 Inductance of a solenoid

What is the inductance of a solenoid (Section 28.6) of length ℓ that has N windings, each of cross-sectional area A, when the current through the device is I?

❶ **GETTING STARTED** I begin by making a sketch of the solenoid (**Figure 29.35**). If I treat the solenoid as being infinitely long, I know from Section 28.6 that the magnetic field is uniform inside the solenoid and zero outside of it. To calculate the inductance of the solenoid, I follow the steps in the Procedure box on the previous page.

Figure 29.35



❷ **DEVISE PLAN** I first need to obtain an expression for the magnetic field magnitude B in terms of the current I. We derived this expression in Section 28.6 and found $B = \mu_0 nI$ (Eq. 28.6), where n is the number of windings per unit length $n = N/\ell$, and so $B = \mu_0 NI/\ell$. Knowing B, I can calculate the magnetic flux, substitute this result into Eq. 29.21, and then solve for the inductance L.

❸ **EXECUTE PLAN** The magnetic flux through one winding is the area A enclosed by the winding multiplied by the magnetic field B inside the solenoid: $\Phi_B = BA$. For the solenoid, the magnetic flux is then

$$\Phi = NBA = \frac{\mu_0 N^2 IA}{\ell}.$$

Substituting this expression for the magnetic flux into Eq. 29.21 yields

$$\frac{d}{dt}\frac{\mu_0 N^2 IA}{\ell} = L\frac{dI}{dt}.$$

The only part of the left side that is time-dependent is the current I, so differentiating with respect to time and simplifying give me

$$L = \frac{\mu_0 N^2 A}{\ell}. \checkmark$$

❹ **EVALUATE RESULT** My result shows that the inductance of a solenoid increases as the square of the number of windings N. This makes sense because both the magnetic field inside the solenoid and the magnetic flux increase with N. That inductance also depends on the area A enclosed by each winding also makes sense because increasing A increases the magnetic flux. Finally, that the inductance is inversely proportional to the length ℓ of the solenoid makes sense because the magnetic field is proportional to the number of turns per unit length.



29.14 A solenoid has 2760 windings of radius 50 mm and is 0.60 m long. If the current through the solenoid is increasing at a rate of 0.10 A/s, what is the magnitude of the induced emf?

29.8 Magnetic energy

We saw in Section 26.6 that work must be done to charge a capacitor and that this work increases the electric potential energy stored in the electric field of the capacitor. Likewise, work must be done on an inductor to establish the current through it because the change in current causes an induced emf that opposes this change; this work increases the **magnetic potential energy** stored in the magnetic field of the inductor.

In Chapter 26, we saw that electric potential energy can be attributed either to the configuration of charge in a system or to the electric field of this charge configuration. Likewise we can attribute the potential energy in an inductor either to the current through it or to the magnetic field caused by the current.

How much potential energy is stored in an inductor in which the current is I? To work this out, we calculate the work required to create a current I in an inductor. We begin by writing the work dW done on the inductor when an amount of charge dq moves through it. This work is the negative of the work done by the induced emf of the inductor:

$$dW = -\mathscr{E}_{\text{ind}}\, dq. \tag{29.22}$$

Using the definitions of current and Eq. 29.19, we obtain for the rate at which work is done

$$\frac{dW}{dt} = -\mathscr{E}_{\text{ind}}\frac{dq}{dt} = -\mathscr{E}_{\text{ind}} I = LI\frac{dI}{dt}, \tag{29.23}$$

and so $dW = LI\,dI$. We can then integrate both sides to obtain the work W done on the inductor to create a current I in it:

$$W = \int dW = L\int I\,dI = \tfrac{1}{2}LI^2. \tag{29.24}$$

To determine how much potential energy is stored in the magnetic field of an inductor, let's choose the zero of magnetic potential energy to be when there is no current through the inductor, and therefore no magnetic field. The magnetic potential energy stored in the inductor when there is a current through it is then equal to the work done to increase the current from zero to I:

$$U^B = \tfrac{1}{2}LI^2. \tag{29.25}$$

This is analogous to Eq. 26.4 for the electric potential energy stored in a capacitor, $U^E = \frac{1}{2}CV_{\text{cap}}^2$, with L taking the place of C and I taking the place of V_{cap}. Because this energy in the inductor is stored in the magnetic field, let's express the energy in terms of the magnetic field. We'll do this for the case of a long solenoid, but the result turns out to be generally applicable. Substituting the expression worked out in Example 29.8 for L into Eq. 29.25, we obtain

$$U^B = \tfrac{1}{2}\frac{\mu_0 N^2 A}{\ell}\,I^2, \tag{29.26}$$

where, as in Example 29.8, N is the number of windings, A is the area of each winding of the solenoid, and ℓ is the length of the solenoid.

In Chapter 26, we used the expression for the electric potential energy U^E to arrive at an expression for the energy density in an electric field: $u_E = \frac{1}{2}\epsilon_0 E^2$ (Eq. 26.6). Let us now use Eq. 29.26 to obtain an expression for the energy density in the magnetic field. Equation 28.6, $B = \mu_0 nI$, where $n = N/\ell$ is the number of windings per unit length, gives us the magnitude of the magnetic field inside a solenoid, and taking the square of this equation yields

$$B^2 = \frac{\mu_0^2 N^2 I^2}{\ell^2}. \tag{29.27}$$

Multiplying the right side of Eq. 29.26 by the factor $\mu_0\ell/\mu_0\ell$ gives

$$U^B = \tfrac{1}{2}\frac{\mu_0 N^2 A I^2}{\ell}\frac{\mu_0\ell}{\mu_0\ell} = \tfrac{1}{2}\frac{\mu_0^2 N^2 I^2}{\mu_0\ell^2}A\ell = \tfrac{1}{2}\frac{B^2}{\mu_0}A\ell. \tag{29.28}$$

Because $A\ell$ is the volume of the region of magnetic field inside the solenoid, dividing the energy U^B by this volume gives us the **energy density** of the magnetic field:

$$u_B \equiv \tfrac{1}{2}\frac{B^2}{\mu_0}. \tag{29.29}$$

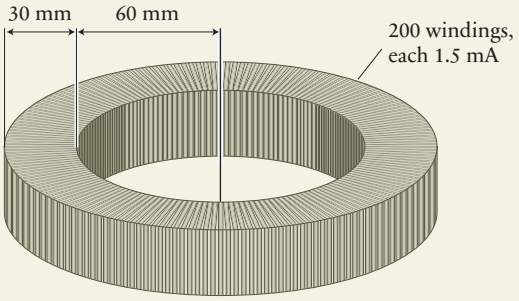
This expression is analogous to the expression for the energy density u_E in the electric field (Eq. 26.6), with B appearing instead of E and $1/\mu_0$ replacing ϵ_0. If the magnetic field is nonuniform, we must subdivide the volume of interest into small enough segments that B can be considered uniform within each segment, then apply Eq. 29.29 to each segment and take the sum of all the contributions. This amounts to integrating the energy density of the magnetic field over the volume containing the magnetic field:

$$U^B = \int u_B\,dV. \tag{29.30}$$

EXAMPLE 29.9 Magnetic energy stored in a square toroid

Consider a toroid with square windings (**Figure 29.36**). The inner radius is $R = 60$ mm, each winding has width $w = 30$ mm, and there are 200 windings, each carrying a current of 1.5 mA. What is the magnetic potential energy stored in this toroid?

Figure 29.36 Example 29.9.



❶ **GETTING STARTED** The current through the toroid causes a magnetic field inside the toroid. This magnetic field stores magnetic potential energy. I note that the width of the windings is not negligible relative to the toroid radius, and this tells me that I must assume a nonuniform magnetic field magnitude across the width of the windings. I therefore must use Eq. 29.30 to determine the amount of magnetic potential energy stored.

❷ **DEVISE PLAN** Before I can use Eq. 29.30, I must obtain the energy density of the magnetic field, which is given by Eq. 29.29. To calculate this energy density, I need to know the magnitude B of the magnetic field. Because the magnetic field is nonuniform, I must determine B by using Eq. 28.9, $B = \mu_0 NI/2\pi r$, where r is the radial distance from the center of the ring formed by the toroid. So to determine the magnetic potential energy stored, I must integrate Eq. 29.29 over the volume of the space enclosed by the windings of the toroid.

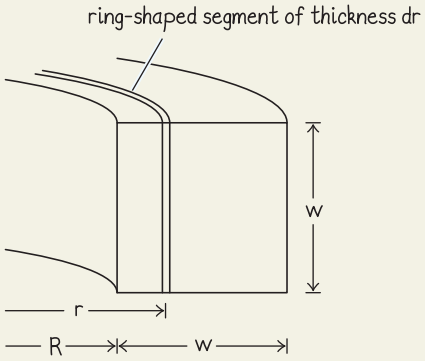
❸ **EXECUTE PLAN** Equation 29.29 gives me

$$U^B = \int u_B\,dV = \frac{1}{2\mu_0}\int B^2 dV = \frac{\mu_0 N^2 I^2}{8\pi^2}\int \frac{dV}{r^2}. \qquad (1)$$

To integrate over the volume of the toroid, I divide the volume into ring-shaped segments (**Figure 29.37**) of height w, radius r, and thickness dr. The volume dV of each segment is equal to the product of the ring's circumference $2\pi r$, height w, and thickness dr: $dV = 2\pi r w\,dr$. Substituting this expression into Eq. 1 and integrating over r from the inner radius of the toroid, $r = R$, to the outer radius, $r = R + w$, I get

$$U^B = \frac{\mu_0 N^2 I^2}{8\pi^2}\int_R^{R+w}\frac{2\pi w r\,dr}{r^2} = \frac{\mu_0 N^2 I^2 w}{4\pi}\int_R^{R+w}\frac{dr}{r}.$$

Figure 29.37



Working out the integration over r yields

$$U^B = \frac{\mu_0 N^2 I^2 w}{4\pi}\ln\left[1 + \frac{w}{R}\right]. \qquad (2)$$

Substituting in the values for μ_0, N, I, R, and w, I obtain for the magnetic potential energy stored in the toroid

$$U^B = (10^{-7}\,\mathrm{T\cdot m/A})(200)^2(1.5\times10^{-3}\mathrm{A})^2 \times (30\cdot10^{-3}\,\mathrm{m})\ln\left(1 + \frac{30\ \mathrm{mm}}{60\ \mathrm{mm}}\right)$$

$$= 1.1\times10^{-10}\,\mathrm{T\cdot m^2\cdot A} = 1.1\times10^{-10}\,\mathrm{J},\ \checkmark$$

where I have used $1\ \mathrm{T\cdot m^2\cdot A} = 1\ [\mathrm{N/(A\cdot m)]\cdot m^2\cdot A} = 1\ \mathrm{N\cdot m} = 1\ \mathrm{J}$ (Eq. 27.6).

❹ **EVALUATE RESULT** The result I obtained is a very small amount of energy, but I have no idea (yet) how much energy is stored in a magnetic field. I can, however, compare my result with the result I would obtain in the limit where the width of each winding is much less than the toroid radius, $w \ll R$. In that limit the magnetic field is uniform inside the toroid, and its magnitude is given by Eq. 28.9, $B = \mu_0 NI/2\pi R$. The volume inside the windings of the toroid is equal to the toroid circumference, $2\pi R$, times the area of a winding, w^2: $V = 2\pi R w^2$. Substituting this expression for B in Eq. 29.29 and multiplying the energy density of the magnetic field by the volume V, I then get for the magnetic potential energy stored in the toroid

$$U^B = \tfrac{1}{2}\frac{\mu_0^2 N^2 I^2/(2\pi R)^2}{\mu_0}(2\pi R w^2) = \frac{\mu_0 N^2 I^2 w^2}{4\pi R}$$

$$= \frac{\mu_0 N^2 I^2 w}{4\pi}\frac{w}{R}. \qquad (3)$$

Because $\ln(1+\epsilon) \approx \epsilon$ for $\epsilon \ll 1$, Eq. 2 is equal to Eq. 3 in the limit that $w \ll R$, giving me confidence in my integration result.

29.15 How does the energy density of a 1.0-T magnetic field compare with the energy density of an 1.0-V/m electric field?

QUANTITATIVE TOOLS

Chapter Summary

Induced emf and current (Sections 29.1, 29.2, 29.4, 29.5)

Concepts A changing magnetic flux through the area enclosed by a conducting loop causes an **induced current** through the loop. This process is called **electromagnetic induction.**

According to **Lenz's law**, the direction of an induced current is always such that the magnetic flux produced by the induced current opposes the change in the magnetic flux that induces the current.

Quantitative Tools If a straight rod of length ℓ moves with speed v perpendicular to a magnetic field of magnitude B, the magnitude of the **induced emf** across the ends of the rod is

$$|\mathscr{E}_{\text{ind}}| = B\ell v. \tag{29.3}$$

According to **Faraday's law,** the induced emf in a closed path due to a changing magnetic flux through the area enclosed by the path is

$$\mathscr{E}_{\text{ind}} = -\frac{d\Phi_B}{dt}, \tag{29.8}$$

where $d\Phi_B/dt$ is the rate at which the magnetic flux changes. The **induced current** in a conducting loop of resistance R is

$$I_{\text{ind}} = \frac{\mathscr{E}_{\text{ind}}}{R}. \tag{29.4}$$

Reminder: The **magnetic flux** Φ_B through a surface is

$$\Phi_B \equiv \int \vec{B} \cdot d\vec{A}, \tag{27.10, 29.5}$$

where the integral is taken over the surface. Magnetic flux is measured in **webers** (Wb), where $1\ \text{Wb} \equiv 1\ \text{T} \cdot \text{m}^2$.

The electric field accompanying a changing magnetic field (Sections 29.3, 29.6)

Concepts A changing magnetic field is accompanied by an electric field. Unlike an electrostatic field, this electric field does not necessarily do zero work on a charged particle that travels around a closed path.

Quantitative Tools The integral around any closed path of the electric field that accompanies a changing magnetic field is given by

$$\oint \vec{E} \cdot d\vec{\ell} = -\frac{d\Phi_B}{dt}, \tag{29.17}$$

where the line integral is around a closed path, and the magnetic flux is through a surface bounded by that same closed path.

Inductance (Section 29.7)

Quantitative Tools If an emf $\mathscr{E}_{\text{ind}}$ is induced in a loop when the current through it changes at the rate dI/dt, the **inductance** L of the loop is given by

$$\mathscr{E}_{\text{ind}} = -L\frac{dI}{dt}. \tag{29.19}$$

See the Procedure box on page 1028 for guidance on how to calculate L. The SI unit of inductance is the **henry** (H):

$$1\ \text{H} \equiv 1\ \text{V} \cdot \text{s/A} = 1\ \text{kg} \cdot \text{m}^2/\text{C}^2. \tag{29.20}$$

The inductance of a solenoid of length ℓ and cross-sectional area A with N turns is

$$L = \frac{\mu_0 N^2 A}{\ell}.$$

Magnetic energy (Section 29.8)

Quantitative Tools The **magnetic potential energy** U^B stored in the magnetic field of an inductor that has inductance L and carries a current I is

$$U^B = \tfrac{1}{2}LI^2. \tag{29.25}$$

The **energy density** u_B of the potential energy stored in a magnetic field of magnitude B, defined as the magnetic potential energy per unit volume, is

$$u_B \equiv \tfrac{1}{2}\frac{B^2}{\mu_0}. \tag{29.29}$$

Questions and Problems

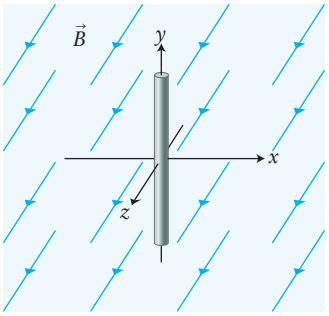
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • *= easy,* •• *= intermediate,* ••• *= hard;* CR *= context-rich problem.*

29.1 Moving conductors in magnetic fields

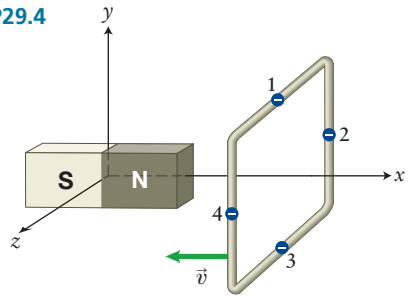
1. What happens if the conducting rod in Figure P29.1 moves (*a*) in either direction along the z axis and (*b*) in either direction along the y axis? •

Figure P29.1



2. The Air Force Thunderbirds aerial demonstration team is performing at an air show located on Earth's magnetic equator. In what directions can the airplanes fly so that there is no charge separation on their metal surfaces? •
3. How should any one of the airplanes described in Problem 2 fly in order to produce the maximum charge separation between its wing tips? (Consider both the direction of motion and the orientation of the wings.) ••
4. Figure P29.4 shows a square conducting loop centered on the x axis, with its sides parallel to the y and z axes. The loop moves with constant velocity in the negative x direction toward a small bar magnet that is centered on the origin and has its polar axis aligned along the x axis. At the instant shown, what is the direction of the magnetic force exerted on an electron sitting at the midpoint of each side? Does the motion of the loop cause any separation of the charge carriers in the loop? Does the motion induce a current in the loop? If so, what is the direction of that current? ••

Figure P29.4

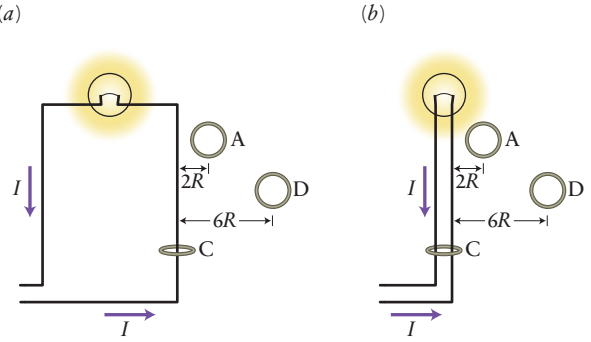


5. What happens if the conducting rod in Figure P29.1 rotates (*a*) clockwise about the z axis as viewed looking down the positive z axis toward the origin, (*b*) clockwise about the x axis as viewed looking down the positive x axis toward the origin, and (*c*) clockwise about the y axis as viewed looking down the positive y axis toward the origin? ••

29.2 Faraday's law

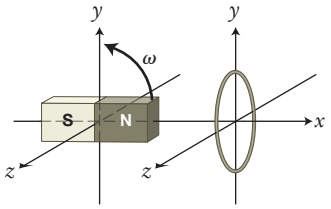
6. The large loop of wire in Figure P29.6*a* carries into and out of a lamp an electric current whose direction alternates back and forth with time. A circular ring of wire of radius R is at one of the three positions labeled A, C, and D. (*a*) At which position is the induced current in the ring greatest? At which position is it smallest? (*b*) Instead of being spread out in large loops, the actual wiring in houses and other buildings has the wires entering and leaving the lamp close and parallel to each other, as shown in Figure P29.6*b*. How is the induced current in the ring affected by this geometry? •

Figure P29.6



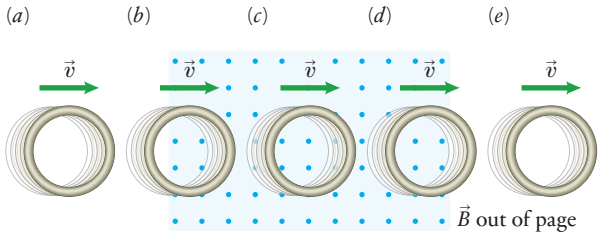
7. The bar magnet in Figure P29.7 is rotated about an axis that runs perpendicular to the page and passes through the center of the magnet. The magnet rotates with rotational speed ω. A loop of wire is placed alongside the magnet as shown. The loop is positioned so that the axis that runs perpendicular to the face of the loop and that passes through the center of the loop lies in the plane in which the magnet rotates. (*a*) Is a current induced in the loop? (*b*) If so, is the induced current constant, or does it vary with time? •

Figure P29.7



8. Suppose that in Example 29.1 you used a circular loop rather than a rectangular loop of the same area. What similarities and differences would you observe in each portion of Figure P29.8 parts *a*–*e*? ••

Figure P29.8



9. A large region of space contains a uniform magnetic field that is increasing with time. You have a piece of wire of length ℓ and want to form a coil from it. What shape should you use for your windings and how many windings should you have to generate the maximum induced current in the coil when it is placed in the changing magnetic field? ••
10. A 1.00-m length of wire is shaped into a rectangle, and another 1.00-m length of identical wire is shaped into a circle. Both wires then move at the same speed into and then out of a region of uniform magnetic field. Is the induced current in the rectangle greater than, smaller than, or the same as the induced current in the circle? •••

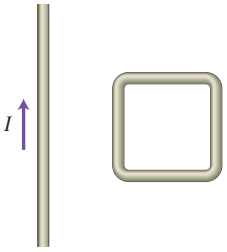
29.3 Electric fields accompany changing magnetic fields

11. You observe a small charge separation between the ends of a conducting rod that is lying on the table in front of you. This is puzzling because the rod is not connected to a battery or other power source; it is at rest on the nonconducting table, pointed toward a nearby window. You call your friend, who calmly asks where you are. You reply that you are riding the high-speed train between Paris and Avignon. Your friend offers explanations for the charge separation in two different inertial reference frames. What are those explanations? •
12. A flat conducting plate is centered in the *xy* plane of an *xyz* coordinate system in which the *xy* plane is horizontal. A bar magnet is lowered vertically toward the plate. If the magnet is oriented so that its north pole faces the origin as the magnet moves toward the plate, what do the electric field lines in the plate look like? What do the field lines look like if the south pole faces the origin as the magnet moves? •
13. A metal pipe is held vertically, and a bar magnet is dropped into it. What do the electric field lines in the pipe look like as the magnet falls through it? ••
14. You have a bar magnet and a circular conducting loop, and you wish to induce a current in the loop that changes direction regularly: clockwise, counterclockwise, clockwise, and so on. Explain how to do this while the loop remains at rest on a wooden table. ••

29.4 Lenz's law

15. You have a bar magnet and a loop of wire. The loop lies in the plane of this page, and we define a positive current as one for which the current direction is clockwise around the loop as viewed from above. How can you move the magnet so as to induce a positive current in the loop? (There are several answers to this question; provide as many as you can.) •
16. If you drop a magnet through a length ℓ of copper pipe, the time interval needed for the magnet to travel through the pipe (even if the magnet never touches the sides of the pipe) is much greater than the time interval the magnet takes to drop the same distance ℓ in air. Why? •
17. The long straight wire in Figure P29.17 carries a current I that varies in time as $I = I_0 \sin(\omega t)$, and a loop of wire is held stationary near the straight wire. When is the induced current in the loop clockwise? When is it zero? When is it counterclockwise? ••

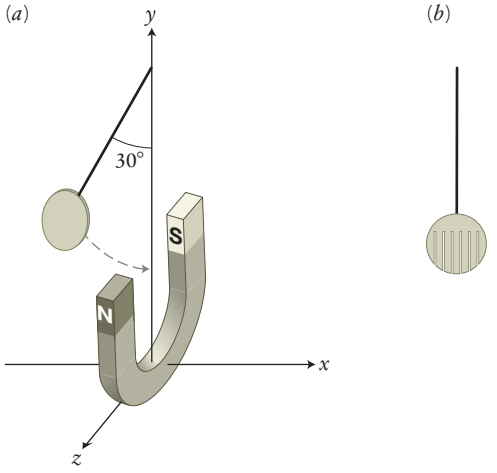
Figure P29.17



18. A pendulum consists of a rod with an aluminum disk attached, suspended from a pivot. First the pendulum is held at a 30° angle to the vertical and released so that it begins to swing. The disk passes between the poles of a powerful magnet on each swing (Figure P29.18*a*). Next a series of vertical slits are cut completely through the disk (Figure P29.18*b*). The pendulum is again held at a 30° angle and released, and the disk again passes between the magnet's poles on each swing. Describe the motion of

the pendulum in each case and explain your answer. Ignore any frictional force exerted at the pivot. ••

Figure P29.18



29.5 Induced emf

19. To measure the magnitude of Earth's magnetic field B_E, you use a single conducting coil with an area $A = 10\ \text{cm}^2$ rotating at an angular speed ω, and measure the peak emf induced on the coil. You know that $B_E \simeq 50\ \mu\text{T}$. What should the minimum value of ω be if you cannot measure emfs below 1 mV? •

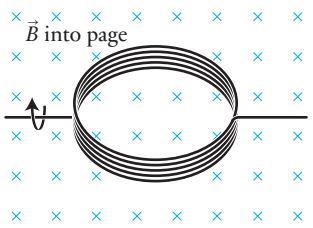
20. A conducting loop having radius 20 cm is held fixed, and there is a magnetic flux of 0.6 T.m^2 through the loop. When the magnetic field is turned off, the magnetic flux through the loop becomes zero in 4 s. What is the average magnitude of the induced emf in the loop during the 4.0 s time interval? •

21. A rectangular loop of length $\ell = 4$ cm, width $w = 3$ cm, and internal resistance $R = 0.5$ V/A is located so that the normal to the loop is parallel to a uniform magnetic field. Suddenly, the loop is turned upside down. A galvanometer measures the total amount of charge flowing through the loop in this time interval to be 1.5 mC. What is the magnitude of the magnetic field? •

22. Suppose you are fabricating a device to measure the orientation of the Erath's magnetic field (i.e., 50 μT). The device can accommodate a single conducting coil attached to a motor that can make it rotate at a speed of 12000 rotations/minute. If the maximum limit of peak induced emf that can ideally be measured by the device is 2×10^{-5} V, then what must be the area A enclosed by the conducting coil that you have to manufacture? •

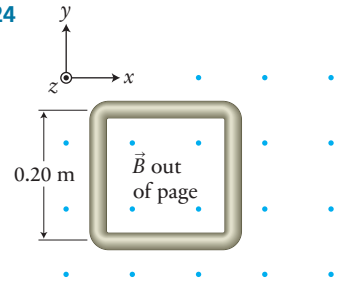
23. A circular coil of radius $R = 50$ mm rotates about an axis that is perpendicular to a uniform magnetic field of magnitude $B = 0.50$ T (Figure P29.23). If the coil completes 60 rotations each second, how many windings must the coil have in order to power an appliance that requires an emf that varies as $\mathcal{E}_{\text{ind}}(t) = V_{\text{max}} \sin(\omega t)$, where $V_{\text{max}} = 155$ V? ••

Figure P29.23



24. The square conducting loop in Figure P29.24 lies in the xy plane of an xyz coordinate system. The loop is in a uniform magnetic field that points in the positive z direction and is decreasing at a rate of 0.070 T/s. What are (*a*) the magnitude of the induced emf in the loop and (*b*) the direction of the induced current in the loop? ••

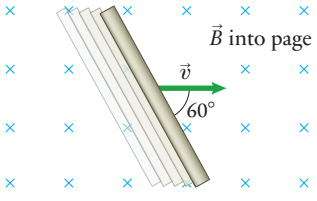
Figure P29.24



25. An increasing number of products, such as passports and credit cards, contain an embedded radio-frequency identification chip that both stores and transmits information. The chips do not have their own power sources. Instead, they receive their power through induction from the device used to read the stored information. The reading device generates a magnetic field, and the passport or credit card containing a coil is passed through this field. If a certain chip needs a peak induced emf of 4.0 V to operate and the magnetic field varies with time as $B(t) = B_{\text{peak}} \sin(\omega t)$, where $B_{\text{peak}} = 5.0$ mT is the maximum magnetic field and $\omega = 8.52 \times 10^7\ \text{s}^{-1}$, what value must the product AN have in order for the chip to operate, where A is the coil's surface area and N is number of turns in the coil? Assume that the magnetic field is perpendicular to the plane of the coil. ••

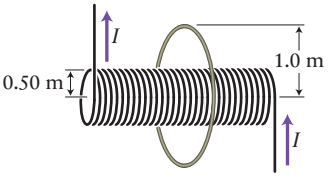
26. A 50-cm-long metal rod is placed in a uniform magnetic field with the rod length perpendicular to the field direction (Figure P29.26). The rod moves at 0.10 m/s, and its velocity vector makes an angle of 60° with the rod length. If the magnitude of the magnetic field is 1.0 T, what is the potential difference between the two ends of the rod? ••

Figure P29.26



27. The very long cylindrical solenoid of Figure P29.27 has a radius of 0.50 m and 1200 windings per meter along its length. A circular conducting loop of radius 1.0 m, and

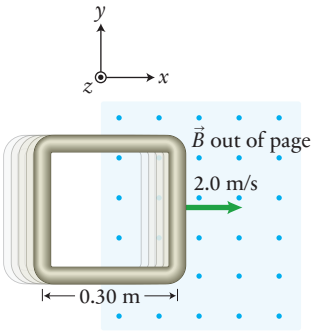
Figure P29.27



resistance 0.5 V/A encircles the solenoid, with the long central axis of the solenoid passing through the center of the loop, and coinciding with the axis of the loop. The solenoid initially carries a steady current $I = 2$ A. If the connection between the solenoid and the battery providing the steady current is interrupted all of a sudden, what will be the total amount of charge flowing in the loop? ●●

28. The space to the right of the y axis in Figure P29.28 contains a uniform magnetic field of unknown magnitude that points in the positive z direction. As a conducting square loop placed in the xy plane (oriented with its horizontal and vertical sides parallel to the x and y axes) moves to the right across the y axis at a constant speed of 2.0 m/s, a 0.24-V emf is induced in the loop. (*a*) If the side length of the loop is 0.30 m, what is the magnitude of the magnetic field? (*b*) What is the direction of the induced current in the loop? ●●

Figure P29.28



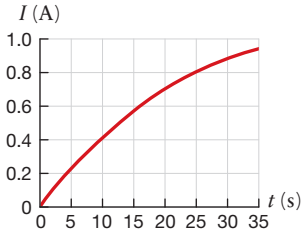
29. A region of space contains a changing magnetic field given by $\vec{B}(t) = B_0 e^{-t/\tau}\hat{k}$, and a circular conducting loop of radius R lies in this region in the xy plane. (*a*) Calculate the upward magnetic flux through the loop as a function of time. Is the flux increasing or decreasing as time passes? (*b*) Calculate the induced emf in the loop as a function of time, where upward is in the $\hat{k}$ direction. (*c*) Determine the current direction in the loop. ●●

30. The coil in a generator has 100 windings and a cross-sectional area of 0.0100 m^2. (*a*) If the coil turns at a constant rotational speed and the magnetic field in the generator is that of Earth ($B = 0.500 \times 10^{-4}$ T), how many 360° rotations must the coil complete each second to generate a maximum induced emf of 1.00 V? (*b*) Based on this calculation, does it seem practical to use Earth's magnetic field in electric generators? ●●

31. You have two cylindrical solenoids, one inside the other, with the two cylinders concentric. The outer solenoid has length $\ell_{\text{outer}} = 400$ mm, radius $R_{\text{outer}} = 50$ mm, and $N_{\text{outer}} = 1000$ windings. The inner solenoid has length $\ell_{\text{inner}} = 40$ mm, radius $R_{\text{inner}} = 20$ mm, and $N_{\text{inner}} = 150$ windings. The inner solenoid is centered within the outer solenoid. When the outer solenoid carries a current given by $I(t) = I_0 \sin(\omega t)$, with $I_0 = 600$ mA and $\omega = 100$ s^{-1}, what is the peak emf in the inner solenoid? ●●

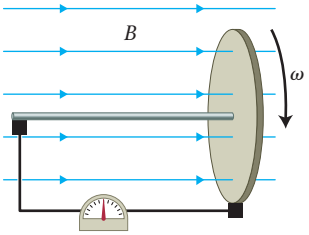
32. A rectangular wire loop enclosing an area $A = 0.40$ m^2 is inside a long cylindrical solenoid with two windings per mm. The area vector of the loop is aligned with the axis of the solenoid, and the current in the solenoid varies according to the graph shown in Figure P29.32. What is the induced emf in the loop at $t = 10$ s? ●●

Figure P29.32



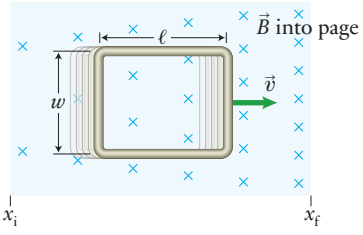
33. A circular metal disk with a shaft through its center rotates about a central axis as shown in Figure P29.33. The unit is placed in a uniform magnetic field of magnitude 1.5 T, directed parallel to the shaft. Two small sliding contacts are placed against the unit, one touching the edge of the disk and the other touching the end of the shaft. (*a*) If the disk radius is 100 mm and the disk completes 300 rotations per minute, what is the value of the emf induced between the two contacts? (*b*) Which contact is at the greater potential? ●●

Figure P29.33



34. The dimensions of the rectangular loop of wire in Figure P29.34 are $\ell = 50$ mm and $w = 40$ mm. The loop moves at speed $v = 30$ mm/s through the magnetic field shown. The field magnitude increases linearly from $B_i(x_i) = 10$ mT at $x_i = 300$ mm to $B_f(x_f) = 60$ mT at $x_f = 900$ mm. The area vector of the loop is parallel to the magnetic field direction. What is the induced emf in the loop? ●●●

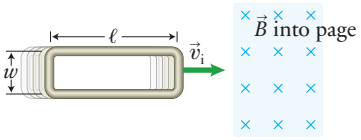
Figure P29.34



35. The dimensions of the rectangular loop of wire in Figure P29.35 are $\ell = 400$ mm and $w = 120$ mm. The mass of the loop is 10 g, its resistance is 5.0 V/A, and before entering the magnetic field shown the loop is moving at a constant speed v_i. The magnetic field magnitude is

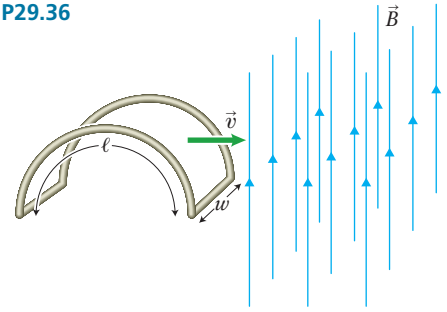
5.0 T, and the loop's area vector is parallel to the field direction. How many seconds after the leading edge of the loop enters the magnetic field is the loop's speed half its initial value? Ignore any friction effects. •••

Figure P29.35



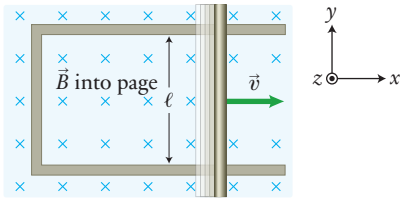
36. (*a*) A rectangular loop of wire is $\ell = 800$ mm long and $w = 500$ mm wide. You bend the ℓ sides into a semicircle while keeping the w sides straight, as shown in Figure P29.36, and then move the curved loop at speed $v = 0.800$ m/s into a uniform magnetic field of magnitude $B = 0.600$ T. One of the sides labeled w is the leading edge of the moving loop, and the magnetic field direction is perpendicular to those sides. What is the value of the emf induced in the loop? (*b*) What is the induced emf if the loop is rotated 90° so that the side leading into the magnetic field is one of the curved sides? •••

Figure P29.36



37. A bar is sliding along a set of connected conducting rails as shown in Figure P29.37. The bar is given an initial velocity $\vec{v}_i$ to the right and then allowed to move freely. The bar has mass m, and the distance between the rails (which is also the bar length) is ℓ. The resistance of the conducting loop comprising the bar and conducting rails is R. Show that the speed of the bar decreases exponentially with time, and determine the time constant. •••

Figure P29.37



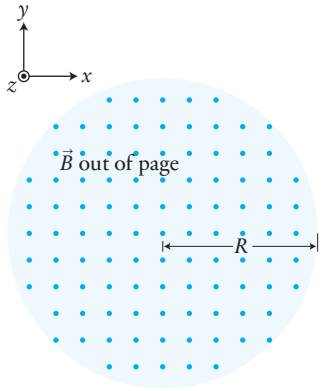
38. The area vector of a rectangular loop that is initially $\ell_i = 300$ mm long and $w_i = 200$ mm wide is aligned with the direction of a uniform magnetic field that has initial magnitude $B_i = 500$ mT. You want to increase the field magnitude smoothly to $B_f = 700$ mT in 10 s, but you want to have zero induced current in the loop. You plan to accomplish this by changing the loop's length ℓ as you increase the magnetic field. In order to have $I_{ind} = 0$, what must the loop length be after 3.00 s and after 10.0 s? •••

29.6 Electric field accompanying a changing magnetic field

39. A uniform magnetic field exists in a circular area. A particle carrying charge $q = 3.0\ \mu$C is placed in the field a distance $r_p = 5$ cm from the center of the circular area. If the magnetic field drops at a constant rate of 0.2 T/s, what its the magnitude and direction that an external force should have so that the particle remains fixed in its position? •

The next four problems are based on the situation illustrated in Figure P29.40. A uniform magnetic field pointing in the positive z *direction fills a cylindrical volume of space of radius* R *whose central axis is the* z *axis. Outside this region, there is no magnetic field.*

Figure P29.40



40. In Figure P29.40, $R = 0.12$ m and a changing magnetic field creates an electric field that has magnitude $E = 10$ V/m at a radial distance $r = 0.060$ m from the center of the magnetic field. What is the instantaneous magnitude of the time rate of change of the magnetic field? •

41. If in Figure P29.40 $R = 0.25$ m and the magnetic field magnitude is decreasing at a rate of 0.30 T/s, what is the magnitude of the electric field created by this changing magnetic field (*a*) at radial distance $r = 0.20$ m from the magnetic field center and (*b*) at $r = 0.50$ m from the center? (*c*) Do the electric field lines encircle the z axis clockwise or counterclockwise, when viewed looking down the positive z axis toward the origin? ••

42. If the magnitude of the magnetic field in Figure P29.40 changes with time as $B = B_{max} \sin(\omega t)$, calculate the magnitude of the electric field that accompanies this changing magnetic field as a function of time t and radial distance r from the center of the magnetic field for (*a*) $r < R$ and (*b*) $r > R$. (*c*) At $t = 0$, what is the direction of the electric field? ••

43. The magnitude of the magnetic field in Figure P29.40 changes with time, and the magnitude of the accompanying electric field inside the magnetic field is given by $E(r, t) = 3Crt^2$, where C is a positive constant and r is the radial distance from the center of the magnetic field. For the case where $B = 0$ at $t = 0$, derive an expression for the magnitude of the magnetic field as a function of time for $t > 0$. ●●
44. A uniform magnetic field fills a cylindrical volume of radius R, and the field magnitude changes with time at an unknown rate. At a certain instant, the electric field magnitude at a radial distance $r_1 < R$ from the magnetic field center is E_1. At this instant, at what radial distance r from the magnetic field center is the electric field magnitude E equal to $E_1/3$ (*a*) if $r < R$ and (*b*) if $r > R$? ●●
45. It takes an electric field magnitude of about 10^6 V/m to ionize atoms in the air and produce a spark. Is it possible to increase the magnetic field inside a long solenoid quickly enough to generate a spark inside? Take the solenoid radius to be 30 mm. ●●●

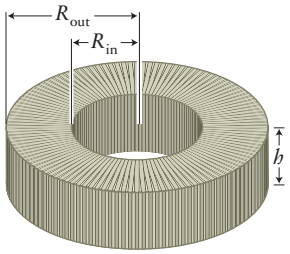
29.7 Inductance

46. An emf of −0.75 V is induced in an inductor of 1.25 H. Find the rate of change of current. Suggest if the current increases or decreases through the inductor. ●
47. If the magnitude of the emf produced in an inductor is 4.0 V when the current through it decreased at a constant rate of 0.2 A/s, calculate the inductance of the inductor. ●
48. A friend of yours needs an inductance of 2 H for one of the circuits he is building. You offer him a cylindrical piece of insulating material of length 30 cm and radius 2 cm to use as a core. How many loops of wire will he need to wrap around this core? ●
49. The current through an inductor of inductance L is given by $I(t) = I_{max} \sin(\omega t)$. (*a*) Derive an expression for the induced emf in the inductor as a function of time. (*b*) At $t = 0$, is the current through the inductor increasing or decreasing? (*c*) At $t = 0$, is the induced emf opposing or aiding the flow of the charge carriers? (Remember that the direction of a positive induced emf is the same as the current direction and the direction of a negative induced emf is opposite the current direction.) (*d*) How are the answers to parts *b* and *c* consistent with the behavior of inductors discussed in the text? ●●
50. The induced emf in an inductor of inductance L varies with time according to $\mathscr{E}_{ind}(t) = -2Ct$, where C is a positive constant. (*a*) If there is no current through the inductor at $t = 0$, calculate the current as a function of time for $t > 0$. (*b*) Is the current increasing or decreasing when $t > 0$? (*c*) Discuss how your answer to part *b* is consistent with the sign of $\mathscr{E}_{ind}$ and what you know about the behavior of inductors. ●●
51. You and a friend have a 0.65-m length of copper wire that has a diameter of 4.115 mm and a wooden rod that is 85 mm long and has a diameter of 10 mm. You're aiming to wind the wire around the rod about 40 times to form an inductor. (*a*) What is the maximum inductance you can achieve with this device? (*b*) Your friend solves this problem by reasoning that a cylindrically wound wire whose length greatly exceeds its radius is a solenoid, and therefore he can use the equation from Example 29.8 to determine the inductance:

$$L = \frac{\mu_0 N^2 A}{\ell}$$
$$= \frac{(4\pi \times 10^{-7}\ \text{T}\cdot\text{m/A})(40^2)\pi(0.0020575\ \text{m})^2}{0.65\ \text{m}}$$
$$= 4.1 \times 10^{-8}\ \text{H}.$$

Evaluate your friend's work. ●●
52. The toroid in Figure P29.52 has 200 rectangular windings, and the toroid radii are $R_{in} = 160$ mm and $R_{out} = 240$ mm. The height of each winding is $h = 20$ mm, such that the rectangular cross section of each winding is $(R_{out} - R_{in}) \times h = (80\ \text{mm}) \times (20\ \text{mm})$. What is the inductance of the toroid? ●●●

Figure P29.52



29.8 Magnetic energy

53. Estimate the amount of magnetic potential energy stored in a 1.50 H inductor when the current through it is 3.0 A. ●
54. If 24 J of magnetic potential energy is stored in a 3.0 H inductor, calculate the current in the inductor. ●
55. A cylindrical solenoid, of radius 0.5 cm and length 30 cm has 1200 windings and carries a current of 0.5 A. (*a*) What is the inductance of the solenoid? (*b*) What is the total magnetic potential energy inside the solenoid? (*c*) What is the magnetic energy density inside the solenoid? ●●
56. When you unplug a coffee maker that plugs into the utility outlet of your car, you notice a spark. Worried that the coffee maker might be broken, you take it apart and find that the heating element is a tungsten wire that, according to the label, is wound 600 times around what appears to be a nonmagnetic core. The core has radius $R = 2.0$ mm and length $\ell = 200$ mm. On the outside of the coffee maker, it says 12 V, 3.2 A. Where does the spark come from? ●●
57. The magnitude of the magnetic field in a magnetic resonance imaging (MRI) machine can be as great as $B = 3.0$ T. Under normal circumstances, this field cannot be shut off by just flipping a switch. Instead the magnitude needs to be carefully decreased to zero. In an emergency, however, the magnet can be "quenched" so that B reduces to zero in 20 s. Such a quench can cost thousands of dollars and likely damages the magnets. Assume that the magnetic field exists inside a cylinder of radius $R = 300$ mm and length $\ell = 200$ mm. How much magnetic potential energy is dissipated when the

magnetic field is quenched in this way, and what is the average rate at which energy is dissipated? ••

58. What is the magnetic potential energy stored in a cylindrical volume of height $h_{\text{cylin}} = 50$ mm and radius $R_{\text{cylin}} = 24$ mm that symmetrically surrounds an infinitely long wire that has radius $R_{\text{wire}} = 2.1$ mm and carries current $I = 2.3$ A? •••

59. A cylindrical solenoid of length ℓ and radius R has n windings per unit length and carries a current I. (*a*) Use the inductance expression $L = (\mu_0 N^2 A)/\ell$ from Example 29.8 and Eq. 29.25, $U^B = LI^2/2$, to derive an expression for the magnetic potential energy in the solenoid as a function of the current and the solenoid's dimensions. (*b*) Derive this same expression by beginning with Eq. 29.30, $U^B = \int u_B \, dV$. •••

Additional Problems

60. Calculate the length of the cylindrical solenoid that has 900 windings, such that the radius of each winding is 2 cm, that shows an inductance of 8.5×10^{-2} H. •

61. The current in a cylindrical solenoid is increased smoothly from $I_i = 0.40$ A at $t_i = 3.0$ s to $I_f = 1.20$ A at $t_f = 5.0$ s. What is the induced emf during this time interval if the solenoid length is $\ell = 150$ mm, the radius of each winding is $R = 20$ mm, and there are 400 windings? •

62. A uniform magnetic field exists in a cuboidal volume of space with $l = 3$ cm, $b = 2$ cm, and $h = 1$ cm. If the magnetic energy stored in the volume is 10 J, what is the magnitude of the magnetic field? •

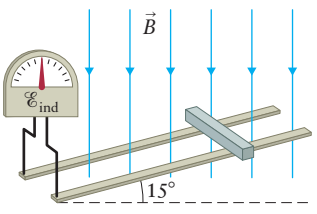
63. A cylindrical electromagnet produces a uniform 1.5-T magnetic field between the poles. If the north and south poles of the electromagnet have circular cross-section having radius 60 mm and separation 50 mm, then calculate the magnetic potential energy stored in the magnetic field of the electromagnet. •

64. The rate at which the magnetic flux inside a wire loop decays is given by $\Phi_B(t) = \Phi_{B,i} e^{-\beta t}$, where $\beta = 0.50/\text{s}$ and $\Phi_{B,i} = 4.0$ Wb. What is the magnitude of the induced current at $t = 10$ s if the resistance of the wire is $R = 2.0$ V/A? ••

65. To construct a solenoid having cross-sectional area 1.256×10^{-3} m^2 and 500 windings that can produce a magnetic field of 10×10^{-3} T, (*a*) calculate the length of the solenoid if the current in the device is 5 A. Calculate (*b*) the inductance of the solenoid, (*c*) considering the cross-sectional area and given magnetic field, the amount of magnetic potential energy stored in the solenoid. (*d*) Using the value of L obtained in part (*b*) estimate the magnetic potential energy stored within the solenoid. ••

66. A conducting bar that is $\ell = 40$ mm long can slide with negligible friction on two parallel conducting rails positioned at an incline of $\theta = 15°$ (Figure P29.66). Initially, the bar is at rest in a magnetic field of magnitude $B = 0.60$ T. What is the emf between the rails 0.20 s after the bar starts sliding while it is still inside the magnetic field? ••

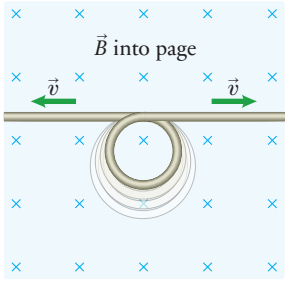
Figure P29.66



67. You have a circular wire loop of radius $a = 20$ cm. It carries a current that increases linearly from 0 to 5 A in 0.01 s. At the center of this loop is a wire loop of radius $b = 1$ mm and resistance $R = 1.8$ mV/A. The loops are coplanar. What is the current induced in the smaller loop? ••

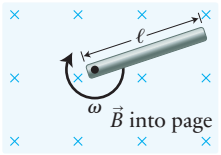
68. You shape a flexible wire into a loop of initial radius $r_i = 30$ mm. You then place the loop in a uniform magnetic field and pull the two ends of the wire in opposite directions at $v = 6.0$ mm/s so that the loop starts closing up (Figure P29.68). If the field magnitude is $B = 0.70$ T, what is the emf between the ends of the wire 2.0 s after you begin pulling? ••

Figure P29.68



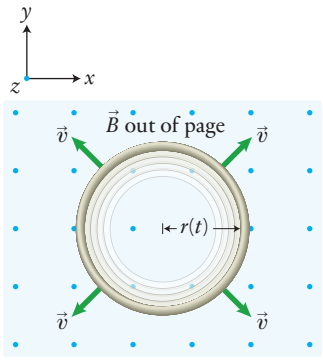
69. A rod of length ℓ rotates about one end at a constant rotational speed ω. The rod is in a uniform magnetic field $\vec{B}$, and the axis of rotation is parallel to $\vec{B}$ (Figure P29.69). (*a*) In terms of ℓ, B, and ω, what potential difference develops between the two ends of the rod? (*b*) Which end of the rod has a greater potential? (*c*) If $\ell = 0.15$ m, $B = 1.0$ T, and $\omega = 377$ s^{-1}, what is the potential-difference value? ••

Figure P29.69



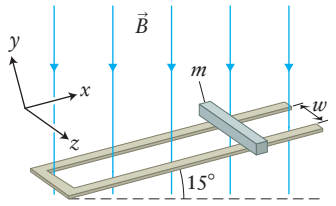
70. A uniform magnetic field of magnitude B fills all space and points in the positive z direction (Figure P29.70). A circular conducting loop in the xy plane is growing larger, with its radius as a function of time given by $r(t) = vt$, where v is a positive constant. (*a*) Derive an expression for the magnitude of the emf induced in the loop as a function of time. (*b*) What is the direction of the induced current? ••

Figure P29.70



71. A conducting bar of width $w = 0.12$ m and mass $m = 8.0$ g can slide freely on two parallel conducting rails positioned at an incline of $\theta = 15°$ (Figure P29.71). The rails are connected at their base by a piece of conducting material. The distance between the rails is w, and the resistance in the rails is $R = 0.20\ \Omega$/A. A uniform magnetic field $\vec{B}$ is exactly vertical as indicated in the drawing, with $B = 0.50$ T. At what constant (terminal) speed does the bar slide down the incline? Ignore friction. ●●

Figure P29.71



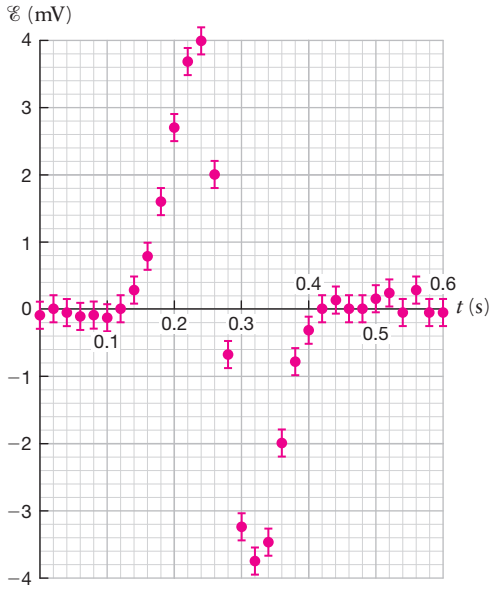
72. For the system described in Problem 71, you measure a constant speed of 1.0 m/s. Because this is not the value you calculated when you ignored friction in the system, you conclude that the friction in the rails is too great to be ignored. What value for the coefficient of kinetic friction μ_k yields the 1.0-m/s speed you measured? ●●●

73. A square loop of wire of side length a and resistance R lies a distance x to the right of a long, straight wire that carries a current I. The straight wire lies in the plane defined by the loop, parallel to one side of the square, and the current direction in the wire is upward. The loop is held stationary, and the current in the wire varies sinusoidally as $I = I_{max} \sin(\omega t)$. Derive an expression for the emf induced in the loop as a function of I_{max}, ω, a, x, and t. ●●●

74. You are working for a company that is interested in mapping geological formations in Earth's crust by examining variations in the magnitude of the magnetic field at Earth's surface. Your assignment is to design a device that can detect small variations in this magnetic field at ground level. Deciding on a system that uses a rotating coil, you realize that the design (number of windings, winding cross-sectional area, rotational speed, and so forth) must give a 1-mV signal in Earth's magnetic field in order to be useful. ●●● CR

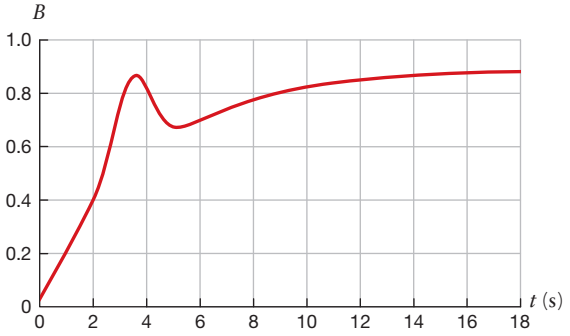
75. On the fifth floor of the physics building, you are in a laboratory class studying induction. You are using a computer to measure emf values from a solenoid when suddenly a thunderstorm breaks out. You see a large blip on the computer screen and 2.0 s later hear a loud thunderclap. Looking out the window, you see that straight ahead a tree was struck by lightning. You quickly save the data on your computer (plotted in Figure P29.75) and note that the solenoid happens to be oriented at an angle of $\theta = 40°$ with respect to your line of sight to the tree. The solenoid windings have cross-sectional area $A = 0.10\ m^2$, and there are $N = 100$ windings. You wonder about the current involved in the lightning flash that struck the tree. ●●● CR

Figure P29.75



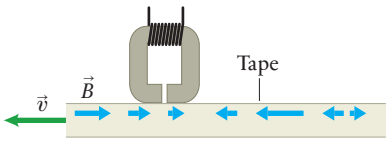
76. An electromagnet in a physics laboratory is damaging electronic apparatus in its vicinity. You suspect that the damage is due to unusually large induced currents created each time the electromagnet power supply is turned on, so you decide to monitor the magnet's magnetic field. Figure P29.76 shows the field magnitude as a function of time during the first 18 s after you connect the magnet to a power supply. Instead of smoothly increasing to its designed value of 0.90 T, the field magnitude starts to decrease after about 3.0 s, very likely due to fluctuations in its power supply. What is not clear is whether these fluctuations can account for the damage, because the electronic apparatus is designed with the expectation that the change in magnetic field strength will be linear with time. ●●● CR

Figure P29.76



77. Analog tape players read sound recorded on magnetic tape. The sound is recorded by magnetizing the tape longitudinally (magnetization parallel to the tape's length) with varying magnetic field magnitudes in either direction. The tape is usually divided into longitudinal tracks so that different signals can be recorded in parallel. In the player, as the tape is pulled across the reading head, a varying emf is induced in a coil wound around the C-shaped metal core of the head (Figure P29.77). This emf is then amplified and sent to the speakers. Trying to learn more about the physics of this technology, you read that the recorded signal magnitude is usually given in terms of *fluxivity*, which is the peak magnetic flux per track width. In particular, you find a paper that gives a fluxivity of 320 nWb/m for a 1000-Hz sine wave recorded on a track 6.3 mm wide [O. Schmidbauer, *J. Audio Eng. Soc.*, 46(10), 859 (1998)]. Naturally, not all of the flux generated threads through the coil, but your search shows that reading heads can have a "flux efficiency" of 98% [J. G. McKnight, B. E. Cortez, and J. A. McKnight, *J. Audio Eng. Soc.*, 46(10), 845 (1998)] and that 250 coil windings are typical. You begin to wonder about the emf induced in the coil and how the frequency of the recorded sound relates to the speed v with which the tape runs. ●●● CR

Figure P29.77





30

Changing Electric Fields

CONCEPTS

30.1 Magnetic fields accompany changing electric fields
30.2 Fields of moving charged particles
30.3 Oscillating dipoles and antennas

QUANTITATIVE TOOLS

30.4 Displacement current
30.5 Maxwell's equations
30.6 Electromagnetic waves
30.7 Electromagnetic energy

As we have seen in Section 29.3, electric fields accompany changing magnetic fields. Is the reverse true, too—do magnetic fields accompany changing electric fields? In this chapter we see that magnetic fields do indeed accompany changing electric fields. Consequently, a changing electric field can never occur without a magnetic field, and a changing magnetic field can never occur without an electric field. The interdependence of changing electric and magnetic fields gives rise to an oscillating form of changing fields called *electromagnetic waves*.

Electromagnetic waves are familiar to us as a wide range of phenomena: visible light, radio waves, and x rays are all electromagnetic waves, the only difference being the frequency of oscillation of the electric and magnetic fields. We see our world by means of these waves, whether by using our eyes to observe our surroundings or by using x-ray diffraction to construct an image of a molecule or a material. Modern communications, from radio and television to mobile telephones, also make extensive use of electromagnetic waves. As we shall see, all these electromagnetic waves consist of changing electric and magnetic fields.

30.1 Magnetic fields accompany changing electric fields

In order to see that a magnetic field accompanies a changing electric field, let's revisit Ampère's law (Section 28.5), which states that the line integral of the magnetic field along a closed path is proportional to the current encircled by the path (Eq. 28.1, $\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I_{\text{enc}}$).

Figure 30.1 shows a current-carrying wire encircled by a closed path. The current encircled by the path is equal to the current through the wire, I. Another way to determine the encircled current is to consider any surface spanning the path and determine the current intercepted by that surface. For example, Figure 30.1 shows two different surfaces spanning the path. The current intercepted by either surface is I, the current encircled by the path.

30.1 Is the current intercepted by the surface equal to the current encircled by the closed path (*a*) in **Figure 30.2*a*** and (*b*) in Figure 30.2*b*?

Figure 30.1 Current-carrying wire encircled by a closed path. Surfaces A and B both span the path. Surface A lies completely in the plane of the path. Surface B extends as a hemisphere whose rim is the path.

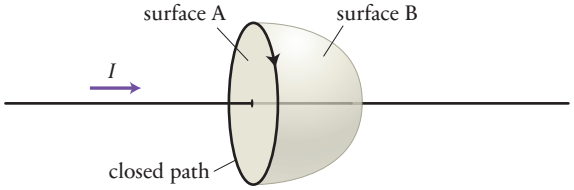
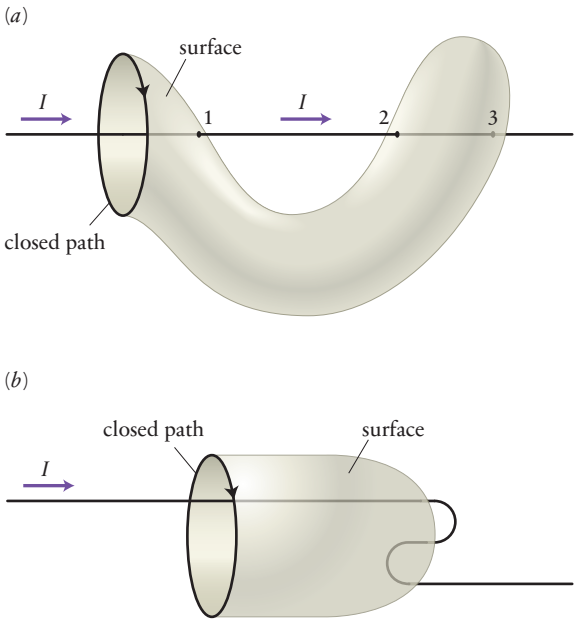


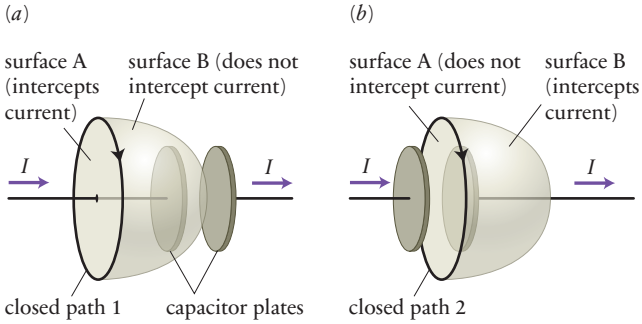
Figure 30.2 Checkpoint 30.1.



Checkpoint 30.1 shows that the current encircled by a closed path is equal to the current that is intercepted by any surface that spans the path, provided we keep track of the directions in which each interception takes place. Ampère's law can equally well be applied to the current encircled by a closed path and to the current intercepted by any surface spanning that closed path.

Now consider inserting an uncharged capacitor into our current-carrying wire while continuing to supply a constant current I to the wire. (That is, the capacitor is being charged.) **Figure 30.3*a*** shows two surfaces A and B spanning the same closed path. The line integral of the magnetic field around the closed path does not depend on the choice of surface spanning the path. However, while the capacitor is charging, surface A is intercepted by a current I but surface B, which passes

Figure 30.3 Capacitor being charged by a current-carrying wire. (*a*) The closed path of interest encircles the wire. Surface A intercepts the current, but surface B passes between the capacitor plates and does not intercept the current. (*b*) The closed path of interest lies between the capacitor plates. Surface A also lies between the plates and does not intercept the current, but surface B intercepts the current.



CONCEPTS

between the capacitor plates, is not intercepted by any current. If we choose a closed path that lies between the capacitor plates (Figure 30.3*b*), a similar difficulty arises. Surface A intercepts no current, while surface B intercepts the current I.

In the case of a capacitor, therefore, the equivalence between encircled current and current intercepted by a surface spanning the encircling path doesn't hold. Surface B in Figure 30.3*a* would lead us to conclude that the line integral of the magnetic field around closed path 1 is zero. Because symmetry requires the magnetic field to always be tangent to the path and have the same magnitude all around the path, the line integral being zero means there is no magnetic field at the location of closed path 1 (even though the path encircles a current). Conversely, surface B in Figure 30.2*b* suggests there is a magnetic field at the location of closed path 2, even though that path encircles no current. Experiments do indeed confirm that there *is* a magnetic field in and around the gap between the plates of the charging capacitor. So only the surfaces that intersect the wires leading to the capacitor appear to provide the correct value of I_{enc} in Ampère's law for both closed paths in Figure 30.3.

Why must there be a magnetic field in and around the gap between the plates of the charging capacitor? Although there is no flow of charged particles between the plates of the capacitor, there *is* an electric field (**Figure 30.4**). Let us examine this electric field in more detail in the next checkpoint.

30.2 (*a*) While the capacitor of Figure 30.4 is being charged, is the current through the wire leading to or from the capacitor zero or nonzero? Is the electric field between the plates zero or nonzero? Is it constant or changing? (*b*) Answer the same questions for the capacitor fully charged.

CONCEPTS

The answers to Checkpoint 30.2 suggest that the magnetic field between the plates of the charging capacitor arises from the *changing* electric field. The current to the capacitor causes the electric field between the plates to change, and the changing electric field between the capacitor plates acts in a way similar to the current that causes this change:

A changing electric field is accompanied by a magnetic field.

When the capacitor is fully charged, the current I into and out of the capacitor is zero, and there is no magnetic field surrounding the wires to the capacitor. Between the capacitor plates, the electric field is no longer changing, and the magnetic field is zero.

There are strong parallels between the electric field that accompanies a changing magnetic field and the magnetic field that accompanies a changing electric field, as **Figure 30.5** illustrates. Experiments show that the electric field lines that accompany a changing magnetic field form loops encircling the magnetic field, just as the magnetic field lines that accompany a changing electric field form loops encircling the electric field.

As we discussed in Section 27.3, the magnetic field surrounding a current-carrying wire forms loops that are clockwise when viewed looking along the direction of the current. The direction of these loops can be described by the right-hand current rule: Point the thumb of your right hand in the direction of the current, and your fingers curl in the direction of the magnetic field. Similarly, the direction of the loops formed

Figure 30.4 Capacitor being charged by a current-carrying wire. The electric field between the plates is shown. Closed path 1 encircles the current through the wire; closed path 2 encircles the electric field between the capacitor plates.

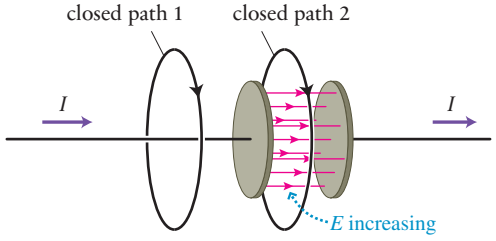
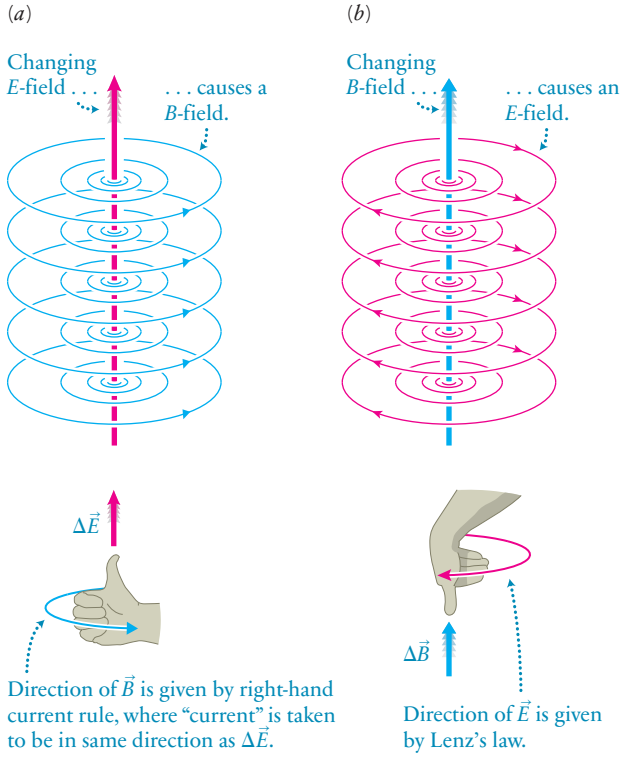


Figure 30.5 Parallels between (*a*) the electric field that accompanies a changing magnetic field and (*b*) the magnetic field that accompanies a changing electric field.



by the magnetic field lines that accompany a changing electric field are given by the right-hand current rule, taking the change in the electric field, $\Delta\vec{E}$, as the "current." If we take this change in the electric field into account in Figure 30.3, treating $\Delta\vec{E}$ as though it were a current, the inconsistency we encountered before is resolved: Either a current or a change in the electric field, $\Delta\vec{E}$, is intercepted by the surface, and so for all surfaces spanning the paths we conclude that there is a magnetic field.



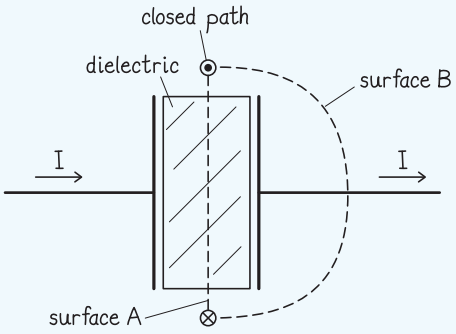
30.3 Consider disconnecting a charged capacitor from its source of current and allowing it to discharge (to release its charge into an external circuit). During discharge, the current reverses direction (relative to its direction when the capacitor was charging), but the electric field between the plates does not change direction. How does the direction of the magnetic field between the plates compare to the direction when the capacitor was charging? Does the right-hand current rule apply?

EXAMPLE 30.1 Capacitor with dielectric

Consider a capacitor being charged with a constant current I and a dielectric between the plates. Is the magnitude of the magnetic field around a closed path spanning the capacitor (such as closed path 2 in Figure 30.4) any different from what it would be without the dielectric? Why or why not?

❶ **GETTING STARTED** I begin by making a two-dimensional sketch of the capacitor, indicating the position of the closed path (**Figure 30.6**).

Figure 30.6



❷ **DEVISE PLAN** To determine the magnetic field magnitude at any position along the closed path, I need to examine the current and the changing electric field intercepted by a surface spanning the closed path.

❸ **EXECUTE PLAN** If I consider a flat surface through the closed path (surface A in Figure 30.6), the surface intersects the dielectric. While the capacitor is charging, the dielectric is being polarized: Negative charge carriers in the dielectric are displaced in one direction, and positive charge carriers are displaced in the opposite direction. This displacement of charge carriers corresponds to a current. Surface A also intercepts a changing electric field. However, without further information about the capacitor, I can determine neither the current nor the electric field between the capacitor plates, which is affected by the presence of the dielectric. Surface A therefore doesn't permit me to compare the magnetic field magnitude to what it would be without the dielectric. I therefore draw another surface, making this surface loop around one of the capacitor plates (surface B in Figure 30.6). This surface intercepts only the wire leading from the capacitor, and I know that the current through the wire is unchanged by the presence of the dielectric. The fact that the current through this wire is unchanged tells me that the effective current in the region containing the dielectric is also unchanged. Therefore, the magnetic field must be the same as it would be without the dielectric. ✔

❹ **EVALUATE RESULT** Intuitively I expect the magnetic field magnitude around my closed path to be unchanged when the magnetic field magnitude around the wires attached to the capacitor is unchanged. The electric field between the capacitor plates gives rise to a displacement of charge carriers within the dielectric and thus affects the electric field between the capacitor plates, but apparently everything adds up to yield, for a given current through the capacitor, the same magnetic field magnitude outside the capacitor for a given current to the capacitor regardless of the presence or absence of the dielectric.

We now have a complete picture of what gives rise to electric and magnetic fields and on what kind of charged particle these fields exert forces. **Table 30.1** summarizes the properties of electric and magnetic fields. Note the remarkable symmetry between the two. Each type of field is produced by charged particles and accompanies a changing field of the other type. Electric fields are produced by charged particles either at rest or in motion, but magnetic fields are produced only by charged particles in motion. Likewise, any charged particle—at rest or in motion—is subject to a force in the presence of an electric field, but only charged particles in motion are subject to forces in a magnetic field.

Table 30.1 Properties of electric and magnetic fields

	Electric field	Magnetic field
associated with	charged particle	moving charged particle
	changing magnetic field	changing electric field
exerts force on	any charged particle	moving charged particle

CONCEPTS

Table 30.2 Electric and magnetic field lines

	Electric field	Magnetic field
lines emanate from or terminate on	charged particle	–
loops encircle	–	moving charged particle
	changing magnetic field	changing electric field

Table 30.2 summarizes what we know about the field lines for electric and magnetic fields. The most striking difference between electric and magnetic fields is that magnetic field lines always form loops but electric field lines do not always form loops. This is a direct consequence of the difference in the sources of these fields. Magnetic field lines must form loops because there is no magnetic equivalent of electrical charge—no magnetic monopole (see Section 27.1). Instead, magnetic fields arise from current loops that act as magnetic dipoles.

Electric and magnetic field lines that represent changing fields both form loops around the changing field. When particles serve as the field sources, however, the difference between magnetic and electric fields is evident: Electric field lines emanate or terminate from charged particles, while magnetic field lines always form loops around moving charged particles (currents).

We shall return to these ideas quantitatively in Section 30.5.

30.4 The neutron is a neutral particle that has a magnetic dipole moment. What does this nonzero magnetic dipole moment tell you about the structure of the neutron?

CONCEPTS

30.2 Fields of moving charged particles

We have seen that capacitors generate changing electric fields when charging or discharging. What else produces changing electric fields? One answer to this question is: changes in the motion of charged particles.

Before examining the electric fields of accelerating charged particles, let's consider the electric fields generated by charged particles moving at constant velocity. **Figure 30.7** shows the electric field of a stationary charged particle and of the same particle moving at constant high speed. (By high speed, I mean a speed near enough the speed of light for relativistic effects to become important.)

The electric field of the stationary particle is spherically symmetrical; the electric field of the moving particle is still radial but definitely not spherically symmetrical. In this electric field, the field lines are sparse near the line along which the particle travels and are clustered together in the plane perpendicular to the motion. (This clustering is a relativistic effect and takes place for the same reason that objects moving at relativistic speeds appear shorter along the direction of motion, as discussed in Sections 14.3 and 14.6.) Consequently, the electric field created by the moving particle is strongest in that perpendicular plane. The faster the particle moves, the more the electric field lines bunch up in the transverse direction.

Keep in mind that as the particle moves at constant speed, the electric field lines move with it. At any instant, the electric field lines point directly away from the position of the particle *at that instant*. This means that as the particle moves, the electric field at a given position changes.

Because the particle in Figure 30.7*b* is moving, it is like a tiny current; it has a magnetic field that forms loops around its direction of travel, as shown in Figure 28.2*b*. The particle in Figure 30.7*a* does not produce a magnetic field because it is at rest.

Now let's consider a particle that is initially at rest and then is suddenly set in motion. The electric field of this particle is shown at three successive instants in **Figure 30.8**. Figure 30.8*b* and *c* show something we have not seen before: electric field lines that do not point directly away from the charged particle that is their source but instead are disrupted by sharp kinks. What is more, these kinks, which appear when the particle accelerates (just after Figure 30.8*a*), do not

Figure 30.7 Electric field line pattern of a charged particle (*a*) at rest and (*b*) moving to the right with speed v (v is a significant fraction of the speed of light). For the moving particle, the electric field lines cluster around the plane perpendicular to the direction of motion.

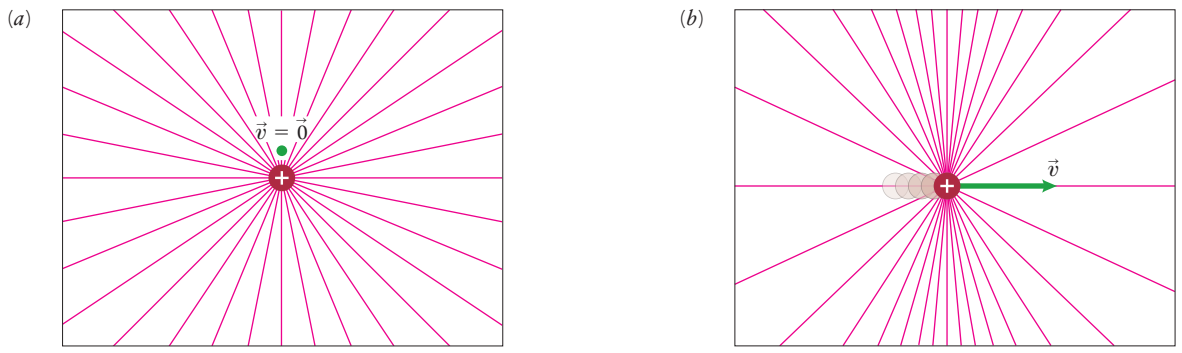
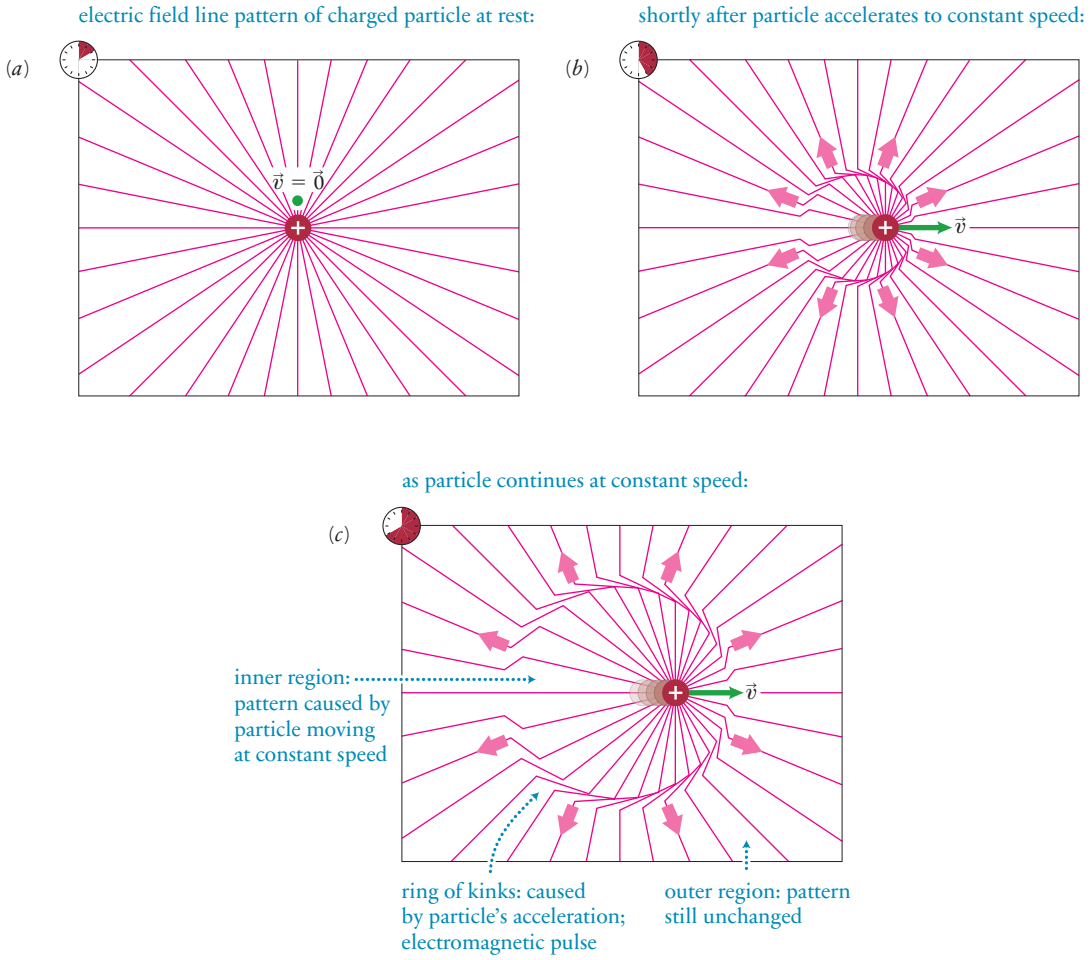


Figure 30.8 Electric field line pattern of a particle (*a*) initially at rest, (*b*) accelerating to speed v, and (*c*) moving at constant speed v. In (*b*) and (*c*), the ring of kinks in the electric field lines traveling outward from the particle corresponds to an electromagnetic wave pulse. Note that the speed v is smaller than the speed of the particle in Figure 30.7, indicated by the shorter arrow. Consequently, the electric field lines here are less sharply bunched around the vertical.



go away once the particle reaches its final constant speed. Instead, they travel radially out from the location where the particle was when it started moving.

Where do these kinks in the electric field pattern come from? They arise because the electric field cannot change instantaneously everywhere in space to reflect changes in the source particle's motion. Remember that field lines extend infinitely far away from the particles that are their sources. If the electric field associated with a particle could change immediately everywhere in the universe when that particle changes its motion, then information about the change in motion would also be transmitted instantaneously throughout the entire universe. As we saw in Chapter 14, however, experiments show that such an instantaneous transmission of information does not happen. Changes in the electric field, and the information that these changes carry, travel at a finite (though very great) constant speed. In fact, in vacuum such changes always travel at the same speed regardless of the details of the motion of the particles that produce them.

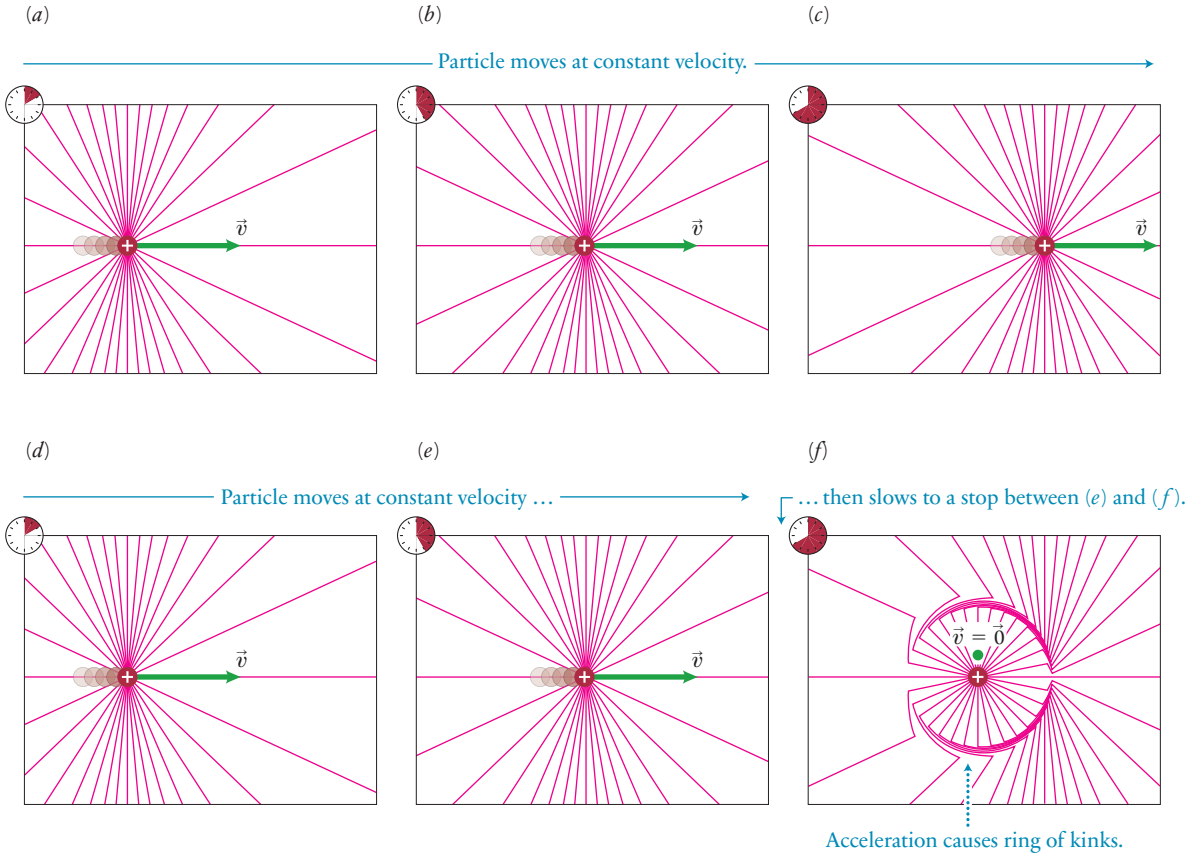
At distances that are too great for changes to reach in the time interval represented in Figure 30.8, the electric field line patterns in Figures 30.8*b* and *c* are still the same as the pattern of the stationary particle of Figure 30.8*a*. At distances that can be reached in that time interval, the electric field line patterns in parts *b* and *c* are those of the moving particle. Kinks form in order to connect these two patterns.

These kinks also form when a particle initially moving at constant velocity $\vec{v}$ abruptly comes to a stop (**Figure 30.9** on the next page). The particle stops just after the instant shown in Figure 30.9*e*. Part *f* shows the electric field line pattern of the stationary particle after some time interval has elapsed.

The electric field line density and consequently the magnitude of the electric field are much greater in the kinks than elsewhere. The energy density in the kinks region is therefore greater than the energy density in

CONCEPTS

Figure 30.9 Electric field line pattern of a charged particle moving at some relativistic speed v. The upper diagrams show successive instants as the particle moves at constant velocity. The lower diagrams show the same instants, but the particle slows down to a stop between (*e*) and (*f*).



other parts of the electric field. As the kinks move, they carry energy away from the particle. These kinks (and the energy carried by them) are one of the two parts of *electromagnetic waves*. As you might guess, kinks in magnetic field lines are the other part. Because changing electric fields are accompanied by changing magnetic fields (and vice versa), the two are always found together. An **electromagnetic wave** is thus a combined disturbance in an electric and a magnetic field that is propagating through space. Because a single isolated propagating disturbance is called a wave pulse (see Section 16.1), the kinks that appear in Figures 30.8 and 30.9 are *electromagnetic wave pulses*.

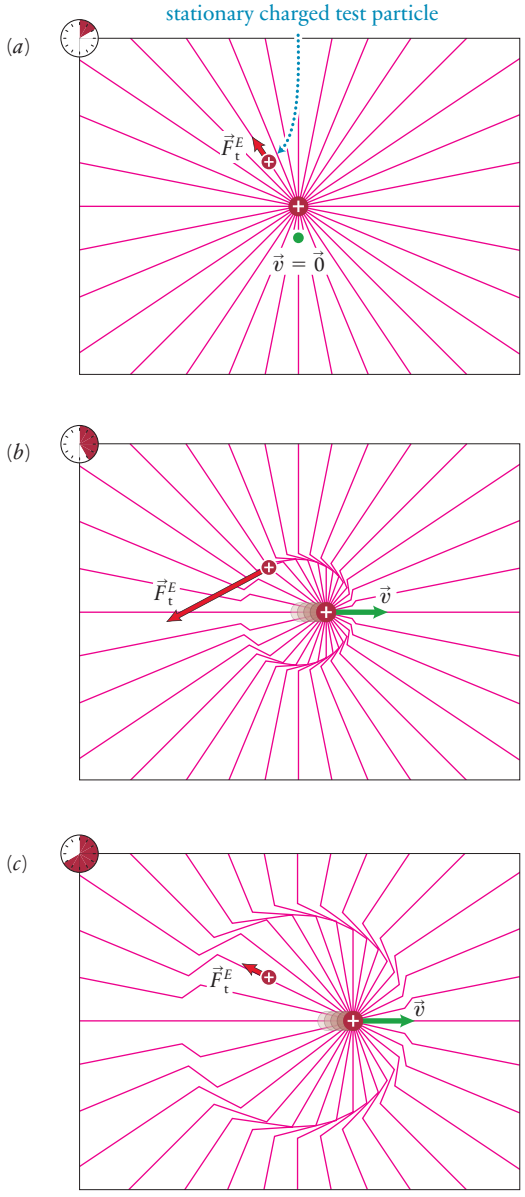
CONCEPTS

30.5 Estimate the final speed v of the charged particle in Figure 30.8 in terms of the speed of propagation c of the electromagnetic wave pulse produced by the particle's acceleration.

Let us now look at what effect an electromagnetic wave pulse has on a charged particle. **Figure 30.10** shows the force exerted on a stationary charged test particle by the electric field of an accelerated charged particle. At the first instant shown (Figure 30.10*a*), before the particle at the center of the panel is accelerated, the force exerted by the electric field on the test particle acts in the direction of the electric field line joining the two particles and points away from the center particle. At the second instant shown (Figure 30.10*b*), the center particle has been accelerated, and the wave pulse created by the acceleration has just reached the test particle. The electric force exerted on the test particle is no longer directed along the line joining the two particles but is directed along the kinks in the electric field lines. The force therefore has a component tangential to a circle centered on the original position of the accelerated particle at the center of the panel. (The exact direction depends on the magnitude and duration of the acceleration of the accelerated particle.) Moreover, because the electric field line density is large in the region of the kinks, the electric force is large in magnitude.

At the final instant shown (Figure 30.10*c*), the wave pulse has traveled beyond the test particle and the electric force once again points away from the particle. The electric field line density is much smaller again, and so the magnitude of the electric force exerted on the test particle is again much smaller.

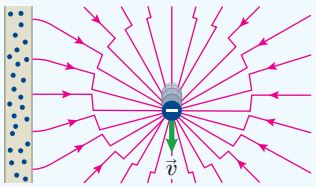
Figure 30.10 Electric force exerted on a stationary charged test particle by the electric field of an accelerated charged particle.



EXAMPLE 30.2 Electromagnetic wave pulse

A particle carrying a negative charge is suddenly accelerated in a direction parallel to the long axis of a conducting rod, producing the electric field pattern shown in **Figure 30.11**.

Figure 30.11 Electric field of an accelerated particle near a conducting rod, before the wave pulse reaches the rod. (The electric field lines bend in near the conducting rod due to the rearrangement of charge carriers at the surface of the rod.)



Does the electric field of the negatively charged particle create a current through the rod (*a*) at the instant shown in the figure, before the electromagnetic wave pulse created by the acceleration reaches the rod, and (*b*) at the instant the pulse reaches the rod? If you answer yes in either case, in which direction is the current through the rod?

❶ **GETTING STARTED** A current is created when charge carriers in the rod flow through the rod. For the carriers to flow, a force needs to be exerted on them.

❷ **DEVISE PLAN** To determine whether there is a current through the rod, I must determine if the electric field is oriented in such a way as to cause a flow of charge carriers through the rod. Even though in a metallic rod only electrons are free to move, I can pretend that only positive charge carriers are free to move because as I saw in Section 27.3, my answer is independent of the sign of the mobile charge carriers.

❸ **EXECUTE PLAN** (*a*) No. Before the pulse reaches the rod, the electric field is constant and so the rod is in electrostatic equilibrium. Therefore the electric field magnitude inside the rod is zero, so no charge carriers in the rod flow at the instant shown. ✔

(*b*) Yes. Once the pulse arrives at the rod, the electric field in the rod points downward, accelerating positively charged particles downward and causing a downward current. ✔

❹ **EVALUATE RESULT** Because the particle being accelerated is negatively charged, it makes sense that it pulls positive charge carriers in the rod along (with a delay caused by the time interval it takes the wave pulse to travel to the rod). In practice, electrons in the rod are accelerated upward, but the result is the same as what I describe for positive charge carriers.

30.6 In Figure 30.10, in which regions of space surrounding the accelerating particle does a magnetic field occur?

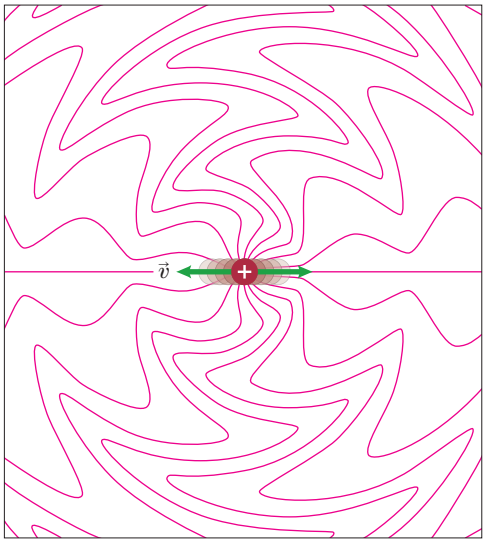
30.3 Oscillating dipoles and antennas

The wave pulse we have just considered is a brief, one-time, propagating disturbance in the electric field, analogous to the disturbance created when the end of a taut rope is suddenly displaced (as in Figure 16.2, for instance). Just as a harmonic wave can be generated on a rope by shaking the end of the rope back and forth in a sinusoidal fashion, a harmonic electromagnetic wave can be generated when a charged particle oscillates sinusoidally. **Figure 30.12** on the next page shows the electric field of a charged particle undergoing sinusoidal oscillation. This electric field consists of periodic kinks traveling away from the particle in a wavelike fashion.

In practice, isolated charged particles are not common. More often, positive and negative charged particles are present together, whether in individual atoms or in solid or liquid materials. Thus, displacing a positive particle leaves a negative particle behind, forming an electric dipole. Let us therefore consider the electric field pattern of an oscillating electric dipole.

CONCEPTS

Figure 30.12 Electric field of a sinusoidally oscillating charged particle.



The electric field pattern of a stationary electric dipole with the positive charged particle above the negative charged particle is shown in **Figure 30.13*a***. What about the electric field pattern of a stationary dipole made up of the same charged particles but with their positions switched, so that the dipole moment $\vec{p}$—which points from the negatively charged end to the positively charged end (see Section 23.4)—has reversed? The corresponding pattern of electric field lines has the same shape as shown in Figure 30.13*a*, but the directions of all the electric field lines are reversed.

Now let's work out the electric field pattern of an oscillating dipole, in which the two particles oscillate back and forth with a period T. We begin by considering the electric field of a dipole that undergoes only a single reversal of its dipole moment (that is, one-half of a single oscillation) rather than oscillating continuously. The dipole starts out as shown in Figure 30.13*a* at instant $t = 0$. The charged particles that constitute the dipole then switch places in a time interval $T/2$ (half a cycle) and remain there.

Figure 30.13*b* shows the electric field pattern at instant $t = T$, after the dipole has been at rest in its new orientation for a time interval $T/2$. We can divide the space surrounding the dipole into the three regions shown. First consider the region sufficiently close to the dipole that the electric field is just the electric field of the stationary dipole in its new orientation. If we denote the speed at which changes in the electric field travel outward by c and the dipole has been stationary for a time interval $T/2$, this innermost region occupies a circle of radius $R = cT/2$. (The origin of our coordinate system is the center of the dipole, midway between the two particles.) Inside this circle, the electric field is that of the stationary dipole, the same shape as shown in Figure 30.13*a* but with the electric field line directions reversed.

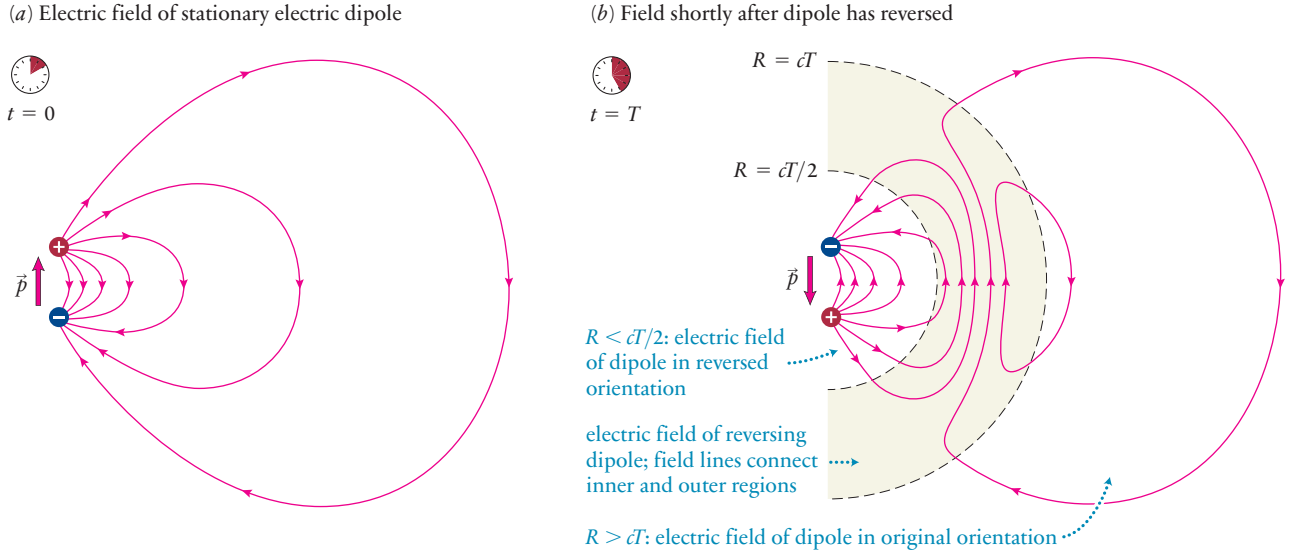
Now consider the region sufficiently far away that no information about the motion of the dipole has reached it yet. This region lies outside a circle of radius $R = cT$. In this region, the electric field pattern is identical to that shown in Figure 30.13*a*, the electric field of the original dipole before it flipped over.

In the highlighted region of Figure 30.13*b* between these two circles, the electric field pattern is not dipolar. Because there are no charged particles in this region, the electric field lines cannot begin or end here. Instead, they must be connected to the electric field lines in the inner and outer regions. Consequently the electric field lines split into two disconnected sets: a set that emanates from the ends of the dipole and a set of loops detached from the dipole.

This electric field line pattern can be generalized to the case of a sinusoidally oscillating dipole (one that doesn't stop after half a cycle). **Figure 30.14** shows snapshots

CONCEPTS

Figure 30.13 (*a*) Electric field of a stationary electric dipole in which the positive particle lies above the negative particle. (*b*) Electric field of the same dipole after the dipole moment has reversed and the charged particles have returned to rest.



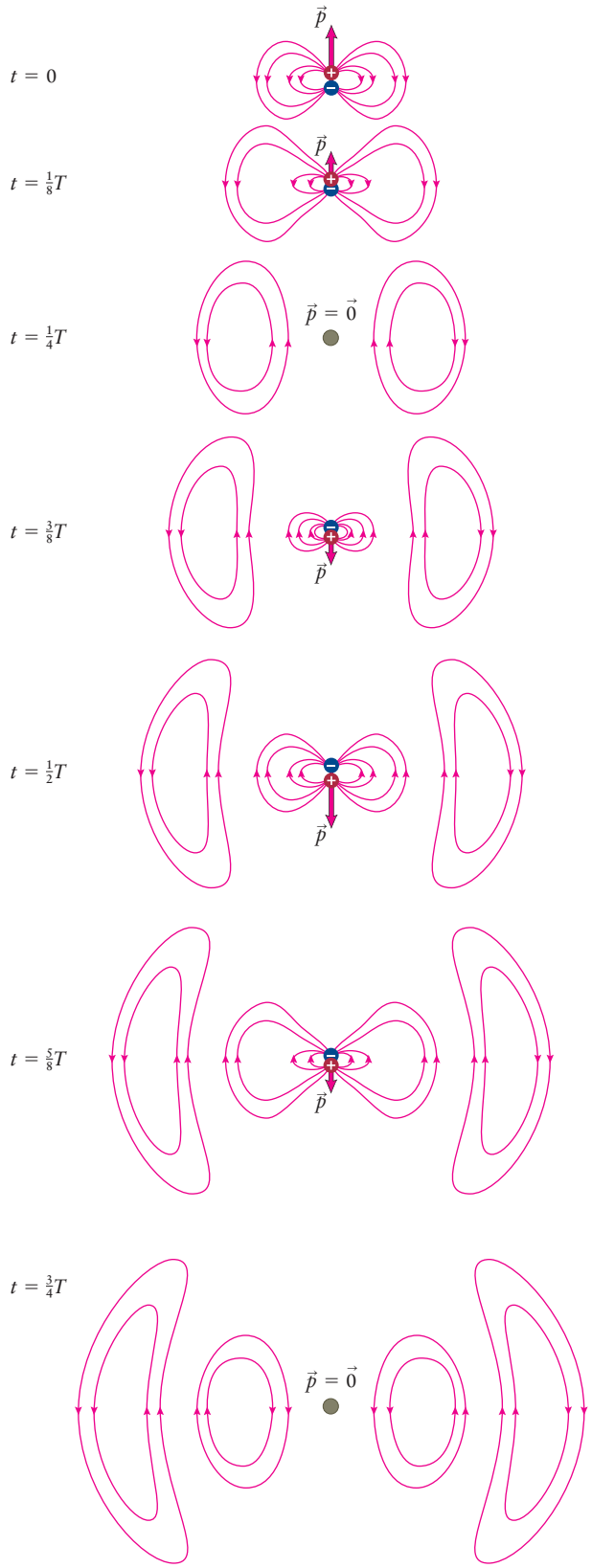


Figure 30.14 Snapshots of the electric field pattern of a sinusoidally oscillating electric dipole at time intervals of $T/8$ (where T is the period of oscillation).

of the electric field pattern of such a dipole at time intervals of $T/8$. Just as with the single half-oscillation, we see a dipolar electric field near the dipole. Farther away, the electric field lines form loops.

Notice how these loops form every half-cycle (that is, at $t = T/4$ and $t = 3T/4$): As the charged particles of the dipole reach the origin during each oscillation, the electric field lines pinch off and the loops travel outward like puffs of smoke. This regular emission of looped electric field lines is a harmonic electromagnetic wave that travels away from the electric dipole horizontally left and right.

30.7 (*a*) If Figure 30.14 shows the oscillating electric field pattern at its actual size, estimate the wavelength of the electromagnetic wave. (*b*) If the wave is traveling at speed $c = 3 \times 10^8$ m/s, what is the wave frequency? (*c*) What is the period of oscillation?

So far we have focused on the electric field pattern of this electromagnetic wave because it is natural to think about the electric field of a dipole. However, the changing electric field of the oscillating dipole is accompanied by a magnetic field. Consequently, the oscillation produces not only an electric field but also a magnetic field.

EXAMPLE 30.3 Magnetic field pattern

Consider the electric field pattern of a sinusoidally oscillating dipole in Figure 30.14. (*a*) At $t = \frac{3}{4}T$, where along the horizontal axis bisecting the straight line connecting the two poles is the electric field increasing with time? Where is it decreasing? (*b*) Based on your answer to part *a*, what pattern of magnetic field lines do you expect in the horizontal plane that bisects the straight line connecting the two poles?

1 GETTING STARTED Because the dipole oscillates sinusoidally, I expect the electric field to be a sinusoidally oscillating outward-traveling wave. The wave is three-dimensional, but the problem asks only about the electric field along the dipole's horizontal axis, so a one-dimensional treatment of this wave suffices. Because the wave is three-dimensional, the amplitude decreases as $1/r$ as the wave travels outward (see Section 17.1), but I'll ignore the decrease over the small distance over which the wave propagates in the figure.

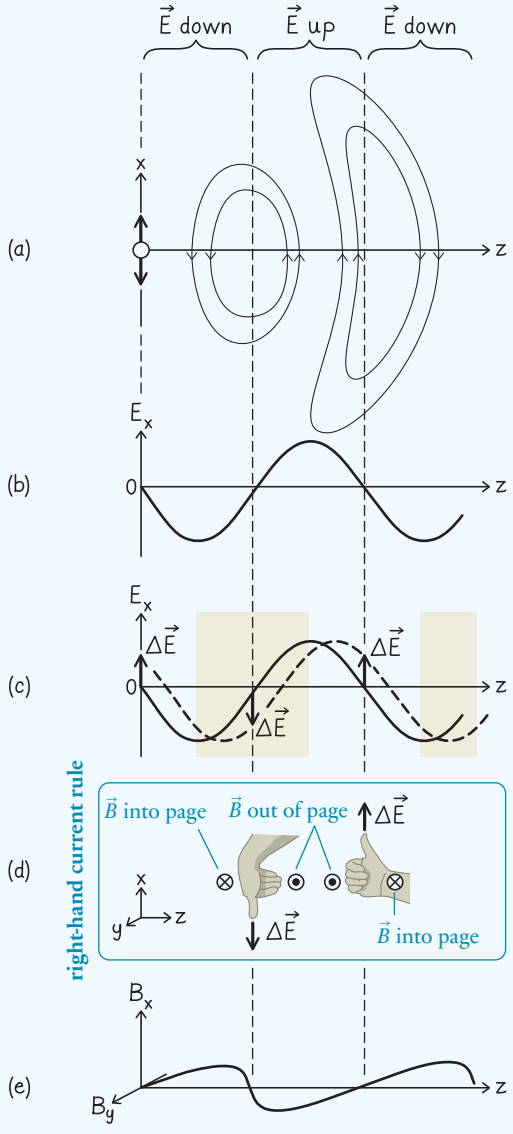
2 DEVISE PLAN I know from Chapter 16 that a one-dimensional sinusoidal wave can be represented by a sine function both in space and in time. (The wave function shows the value of the oscillating quantity as a function of position at a given instant in time, and the displacement curve shows the oscillating quantity as a function of time at a given position.) I can use the information shown in Figure 30.14 to draw the wave function for the electric field at $t = \frac{3}{4}T$. Once I have the wave function and know which way the wave is traveling, I can determine where the electric field increases. Because a changing electric field causes a magnetic field, I can use the information from part *a* to solve part *b*.

(Continued)

CONCEPTS

3 EXECUTE PLAN (*a*) Because the problem asks for information at the instant $t = \frac{3}{4}T$, I begin by copying the right half of the bottom electric field pattern of Figure 30.14 (**Figure 30.15*a***). (The left half is simply the mirror image of the right half.) I draw a rightward-pointing horizontal axis through the center of the dipole and denote this as the z axis. I see that the electric field points downward parallel to the vertical axis (which I take to be the x axis) in the region between the dipole and the center of the first set of electric field loops. In the region between the centers of the first and the second set of loops, the electric field points upward. Because the electric field must vary sinusoidally, I can now sketch how its x component varies with position along the horizontal axis (Figure 30.15*b*).

Figure 30.15

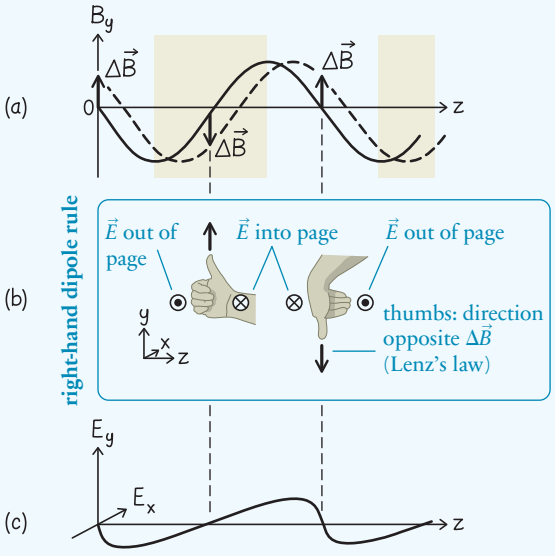


As the wave travels outward, the wave function of Figure 30.15*b* moves to the right (dashed curve in Figure 30.15*c*). The difference between the dashed and solid curves is the change in the electric field $\Delta\vec{E}$ (black arrows); $\Delta\vec{E}$ points down when the derivative of E_x with respect to z is positive (shaded regions) and up when it is negative (unshaded regions). ✔

(*b*) The direction of the magnetic field is determined by the right-hand current rule, taking the direction of the change in the electric field $\Delta\vec{E}$ as the "current." Pointing the thumb of my right hand down in the region where $\Delta\vec{E}$ points down and up where $\Delta\vec{E}$ points up (Figure 30.15*d*), I see from the way my fingers curl that the magnetic field lines form loops in the horizontal (yz) plane that are centered on the vertical black dashed lines, just like the electric field lines do. Consequently the magnetic field points out of the page when E_x is positive and into the page when E_x is negative. If I let the y axis point out of the page, the y component of the magnetic field must be positive when E_x is positive and negative when E_x is negative. I therefore draw a sinusoidally varying function for B_y as a function of position (Figure 30.15*e*). ✔

4 EVALUATE RESULT My answer shows that the electric and magnetic fields have the same dependence on time but are oriented perpendicular to each other. Because the magnetic field changes, Faraday's law tells me that it is accompanied by an electric field. To analyze this electric field I can use an approach similar to the one I used to determine the magnetic field. **Figure 30.16*a*** shows the magnetic wave traveling outward. The difference between the dashed and solid curves is the change in the magnetic field $\Delta\vec{B}$. According to what I learned in Section 29.4, the direction of the electric field accompanying my changing magnetic field is given by Lenz's law and the right-hand dipole rule (Figure 30.16*b*). Consequently the electric field points into the page when B_y is positive and out of the page when B_y is negative. Because the x axis points into the page in this rendering (compare Figures 30.15*d* and 30.16*b*), the x component of the electric field must be positive when B_y is positive and negative when B_y is negative, as shown in Figure 30.16*c*. The electric field shown in Figure 30.16*c* is exactly the electric field I started out with in Figure 30.15*b*. In other words, the electric field yields the magnetic field and the magnetic field yields the electric field, and the two are entirely consistent with one another.

Figure 30.16



CONCEPTS

Figure 30.17 Electric and magnetic field pattern of oscillating dipole. The pink arrows indicate the direction of propagation of the electromagnetic wave pulse. For simplicity, only the fields in the *xz* and *yz* planes are shown.

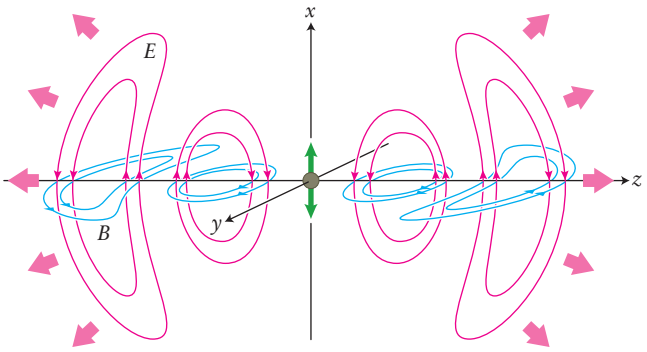
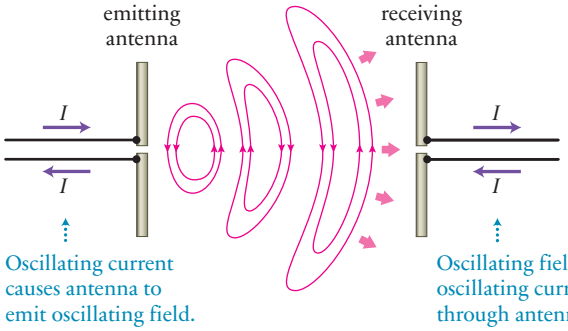


Figure 30.18 System of two antennas, one that emits electromagnetic waves and one that receives them. The emitting antenna is supplied with an oscillating current created by a source of alternating potential difference. An oscillating current is induced in the receiving antenna by the arriving electromagnetic wave.



The solution to Example 30.3 suggests that the magnetic field line pattern is similar to the electric field line pattern, but perpendicular to it. **Figure 30.17** shows the combined electric and magnetic field pattern of an oscillating dipole. Traveling electromagnetic waves, like the one shown in Figure 30.17, are transverse waves (see Section 16.1) because both the magnetic field and the electric field are perpendicular to the direction of propagation. Also, the electric and magnetic fields propagate at the same frequency, and both reach their maxima (or minima) simultaneously; the electric and magnetic fields are therefore in phase with each other.

30.8 (*a*) At the origin of the graphs in Figure 30.15, the electric field is zero, but there is a current due to the motion of the charged particles that constitute the dipole. Is this current upward, downward, or zero at the instant shown in Figure 30.15? (*b*) Is this current (or the absence thereof) consistent with the magnetic field pattern shown in Figure 30.17?

In the wave shown in Figure 30.17, not only are the electric and magnetic fields perpendicular to each other, but throughout the entire wave the electric field has no component perpendicular to the *xz* plane. (The magnetic field, in contrast, is always perpendicular to this plane.) By convention the orientation of the electric field of an electromagnetic wave as seen by an observer looking in the direction of propagation of the wave is called the **polarization** of the wave. An observer looking at the dipole in Figure 30.17 would say that the wave from the dipole is polarized along the *x* axis. Because the electric field oscillation from the dipole in Figure 30.17 retains its orientation as it travels in any given direction, the wave is said to be *linearly polarized*. In certain cases the polarization of an electromagnetic wave rotates as it propagates, and the wave is said to be *circularly* or *elliptically polarized*.

We have seen that oscillating dipoles generate electromagnetic waves by accelerating the oppositely charged particles that make up the dipole in a periodic manner. Practically speaking, how can we cause charged particles, in a dipole or anything else, to accelerate periodically? One common approach is to apply an alternating potential difference to an *antenna*, which is a device that either emits or receives electromagnetic waves. The alternating potential difference drives charge carriers back and forth through the antenna, thereby producing an oscillating current through the antenna.

Antennas that emit electromagnetic waves are designed in many ways to produce a variety of electric and magnetic field patterns. The simplest design is two conducting rods connected to a source of alternating potential difference (**Figure 30.18**). Because of the alternating potential difference, the ends of the antenna are oppositely charged and cycle between being positively charged, neutral, and negatively charged.

When the top end of the antenna is positively charged and the bottom end is negatively charged, the electric field of the antenna points down. When the charge distribution is reversed, the electric field points up. As the charge distribution oscillates, the electric field adjacent to the emitting antenna also oscillates. This changing electric field is accompanied by a changing magnetic field, and the disturbance in the fields travels away from the emitting antenna in the same manner as the electromagnetic wave of Figures 30.14, 30.15, and 30.17.

If the length of each rod in an emitting antenna is exactly one-quarter of the wavelength of the electromagnetic wave emitted, the electric fields produced strongly resemble the dipole fields of Figure 30.17. Such an antenna is often called a *dipole antenna;* it is also called a *half-wave antenna* because the length of the two rods is equal to half a wavelength.

In antennas that receive electromagnetic waves, the oscillating electric field of the wave causes charge carriers in the antenna to oscillate, as discussed in Example 30.3. This produces an oscillating current (shown schematically in Figure 30.18) that can be measured. When operated in this mode, the antenna is said to be *receiving* a signal.

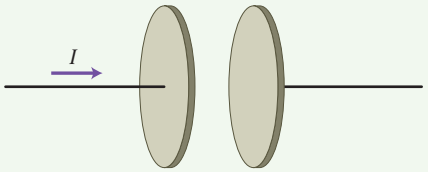
30.9 To maximize the magnitude of the current induced in a receiving antenna, should the antenna be oriented parallel or perpendicular to the polarization of the electromagnetic wave?

CONCEPTS

Self-quiz

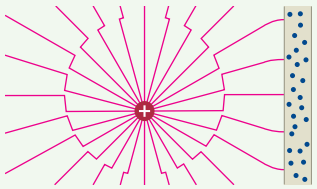
1. Suppose the current shown in **Figure 30.19** discharges the capacitor. What are the directions of $\vec{E}$, $\Delta\vec{E}$, and $\vec{B}$ between the plates of the discharging capacitor?

Figure 30.19



2. A positively charged particle creates the electric field shown in **Figure 30.20**. When the kinks in the electric field lines reach the rod, what is the direction of the current induced in the rod?

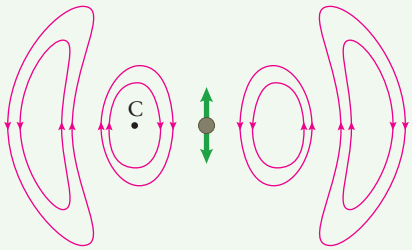
Figure 30.20



3. For the oscillating dipole of Figure 30.14, sketch the electric field pattern at $t = \frac{5}{4}T$.

4. In the electric field pattern for a sinusoidally oscillating dipole shown in **Figure 30.21**, what are (*a*) the direction of the change in the electric field $\Delta\vec{E}$ at point C as the electric field propagates and (*b*) the direction of the magnetic field loop near C?

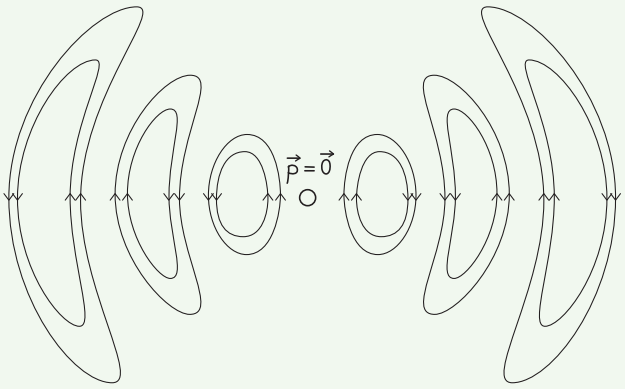
Figure 30.21



Answers

1. The current brings positive charge carriers to the left plate and removes them from the right plate. For the capacitor to discharge, the left plate must be negatively charged and the right one positively charged; this means $\vec{E}$ points left. The electric field decreases as the capacitor discharges, so $\Delta\vec{E}$ is to the right (just as the current is). The magnetic field lines are circular and centered on the axis of the capacitor in a direction given by the right-hand current rule with the thumb along the direction of $\Delta\vec{E}$. That is, the magnetic field lines are clockwise looking along the direction of the current.
2. Because the particle carries a positive charge, the electric field lines radiate outward. The electric field in the kinks therefore points up, and so the kinks induce an upward current through the rod.
3. See **Figure 30.22**. Because the loops move outward and a new pair of loops forms every half-period, the pattern now has three loops. Note in Figure 30.14 that the loops closest to the dipole at $t = \frac{1}{4}T$ and $t = \frac{3}{4}T$ have the same shape but opposite directions. Half a period after $t = \frac{3}{4}T$, at $t = \frac{5}{4}T$, the loops closest to the dipole again have the same shape (and the same direction as at $\frac{1}{4}T$). Likewise, at $t = \frac{5}{4}T$ the second closest loops curl in the direction opposite the direction at $\frac{3}{4}T$.

Figure 30.22



4. (*a*) The loops passing at C travel to the left. At the instant shown, the electric field is close to zero, but as the pattern moves to the left, the electric field lines point downward at C, so $\Delta\vec{E}$ is down. (*b*) The thumb of your right hand aligned in the direction of $\Delta\vec{E}$ makes your fingers curl in the direction of the magnetic field: clockwise viewed into the plane of the page.

30.4 Displacement current

The work we did with Ampère's law in Sections 28.4 and 28.5 dealt only with the magnetic field associated with an electric current. As we saw in Section 30.1, though, magnetic fields also accompany changing electric fields, a phenomenon not covered by our Chapter 28 form of Ampère's law. Let us now see how the quantitative formulation of Ampère's law must be modified to account for the magnetic fields that accompany changing electric fields.

To do this, consider the charging capacitor in **Figure 30.23**. Ampère's law relates the integral of the magnetic field around a closed path to the current intercepted by a surface spanning the path (Eq. 28.1). Applying Ampère's law to surface A in Figure 30.23, we have

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I. \tag{30.1}$$

For surface B, however, the right-hand side of Eq. 30.1 is zero because the current is zero in the gap between the capacitor plates. As we discussed in Section 30.1, there is a change in the electric field between the capacitor plates, so surface B does not intercept a current, but it does intercept a change in electric flux. Let us therefore generalize Ampère's law by adding to the right side a term that depends on this change in electric flux.

We choose this term so that when we apply the generalized version of Ampère's law to the capacitor shown in Figure 30.23, for example, the magnetic field around the designated path is the same whether we calculate it from the current intercepted by surface A or from the change in electric flux $d\Phi_E/dt$ through surface B.

To obtain this generalizing term, let's determine a mathematical relationship between $d\Phi_E/dt$ through surface B and the current to the plates. First, note that the change in electric flux is related to the change in the charge q on the plates, which, in turn, is related to the current I to the plate. Consider the closed surface surrounding the left capacitor plate in Figure 30.23 made up by surfaces A and B combined. Applying Gauss's law to this closed surface, we find that the electric flux through it is

$$\oint_{\text{A+B}} \vec{E} \cdot d\vec{A} = \frac{q}{\epsilon_0}, \tag{30.2}$$

where q is the charge on the capacitor plate. Because the electric field is confined to the region between the plates, the electric flux through surface A is zero (Figure 30.23) and so

$$\oint_{\text{A+B}} \vec{E} \cdot d\vec{A} = \int_{\text{A}} \vec{E} \cdot d\vec{A} + \int_{\text{B}} \vec{E} \cdot d\vec{A} = \int_{\text{B}} \vec{E} \cdot d\vec{A} = \frac{q}{\epsilon_0}. \tag{30.3}$$

If we denote the electric flux through surface B by Φ_E, we see from Eqs. 30.3 that $q = \epsilon_0 \Phi_E$. The rate of change of the charge on the capacitor plates, dq/dt, is equal to the current supplied to the capacitor, so

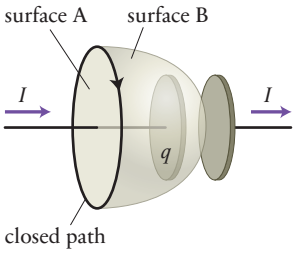
$$I \equiv \frac{dq}{dt} = \epsilon_0 \frac{d\Phi_E}{dt}, \tag{30.4}$$

which is the relationship we were looking for.

If we substitute Eq. 30.4 in the right side of Eq. 30.1, we obtain

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}. \tag{30.5}$$

Figure 30.23 Capacitor being charged by a current-carrying wire. Because surfaces A and B both span the closed path shown, either surface can be used to calculate the magnetic field around the path. Ampère's law must be the same in either case.



We can now use this expression to determine the line integral of the magnetic field around the closed path in Figure 30.23 by evaluating the change in electric flux through surface B. Because the right side of Eq. 30.5 is equal to $\mu_0 I$, we obtain for $\oint \vec{B} \cdot d\vec{\ell}$ the same value we found using the original form of Ampère's law with the current intercepting surface A.

To account for both a current and a changing electric flux, we generalize Ampère's law as follows:

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I_{\text{int}} + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}. \tag{30.6}$$

This equation holds for any surface spanning a closed path and is sometimes called the *Maxwell-Ampère law*, in honor of the Scottish physicist James Clerk Maxwell (1831–1879), who first introduced the additional term in Eq. 30.6. To reflect the fact that we must only include the current intercepted by the surface, not the current encircled by the integration path, we write I_{int} rather than I_{enc}.

The quantity on the right side of Eq. 30.4 is called the **displacement current:**

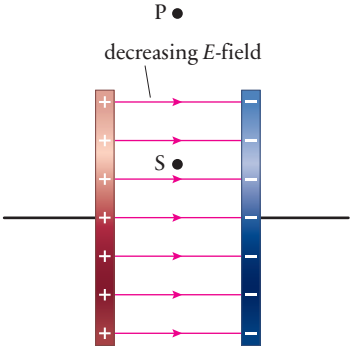
$$I_{\text{disp}} \equiv \epsilon_0 \frac{d\Phi_E}{dt}. \tag{30.7}$$

As you can see from Eq. 30.4, the SI units of the displacement current are indeed those of a current. The name is somewhat misleading because the derivation holds for a capacitor in vacuum where no charged particles are present in the space between the plates. Even if the term is somewhat of a misnomer, it is still useful to associate the change in the electric field with a "current" to determine the direction of the magnetic field accompanying a changing electric field. As we argued in Section 30.1, the direction of this displacement current is the same as that of the change in the electric field $\Delta\vec{E}$. We can then use the right-hand current rule to determine the direction of the magnetic field from the displacement current (see, for example, Figure 30.5).

Using Eq. 30.7, we can write Eq. 30.6 in the form

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 (I_{\text{int}} + I_{\text{disp}}). \tag{30.8}$$

Figure 30.24 Checkpoint 30.10.

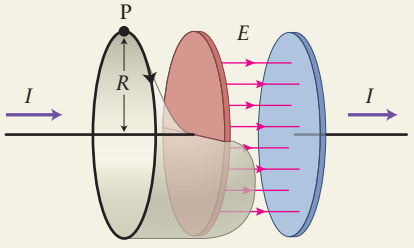


30.10 The parallel-plate capacitor in **Figure 30.24** is discharging so that the electric field between the plates *decreases*. What is the direction of the magnetic field (*a*) at point P above the plates and (*b*) at point S between the plates? Both P and S are on a line perpendicular to the axis of the capacitor.

EXAMPLE 30.4 A bit of both

The parallel-plate capacitor in **Figure 30.25** has circular plates of radius R and is charged with a current of constant magnitude I. The surface is bounded by a circle that passes through point P and is centered on the wire leading to the left plate and perpendicular to that wire. The surface crosses the left plate in the middle so that the top half of the plate is on one side of the surface and the bottom half is on the other side. Use this surface and Eq. 30.6 to determine the magnitude of the magnetic field at point P, which is a distance $r = R$ from the capacitor's horizontal axis.

Figure 30.25 Example 30.4.



QUANTITATIVE TOOLS

❶ **GETTING STARTED** The surface intercepts both a current (through the wire) and a changing electric field (between the plates). To apply Eq. 30.6, I must therefore determine both the current and the electric flux intercepted by the surface.

There also is a simple way to obtain the answer to this question: Ampère's law (Eq. 30.1). If I take the circle through point P centered on the wire and perpendicular to it as the integration path, then the integral on the left side of Eq. 30.1 is equal to the magnitude of the magnetic field times the circumference of the circle: $2\pi RB$. The path encircles the current I, so Ampère's law gives me $2\pi RB = \mu_0 I$ and $B = \mu_0 I/(2\pi R)$. Because the magnetic field magnitude cannot depend on the approach used to calculate it, I should obtain the same result using Eq. 30.6 and the surface in Figure 30.25.

❷ **DEVISE PLAN** The left side of Eq. 30.6 is identical to the left side of Eq. 30.1 and therefore equal to $2\pi RB$. To evaluate the right side of Eq. 30.6, I must determine both the ordinary current and the displacement current intercepted by the surface.

❸ **EXECUTE PLAN** If the surface intercepted the entire electric flux between the plates, the displacement current term on the right side of Eq. 30.6 would be equal to $\mu_0 I$. The surface intercepts only half of the electric flux, however, so the displacement current is

$$\epsilon_0 \frac{d\Phi_E}{dt} = \tfrac{1}{2} I.$$

Next I need to determine how much current is intercepted by the surface. To charge the plate uniformly, the current must carry charge carriers evenly to the two halves of the plate—half of the charge carriers go to the top half of the plate and the other half go to the bottom half of the plate (**Figure 30.26**). Because the top half of the plate is to the right of the surface, the current going to the top half of the plate must cross the surface. The current intercepted by the surface is thus $\frac{1}{2}I$, and the right side of Eq. 30.6 becomes

$$\mu_0 I_{\text{int}} + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt} = \mu_0 \left(\tfrac{1}{2} I\right) + \mu_0 \left(\tfrac{1}{2} I\right) = \mu_0 I. \quad (1)$$

Substituting the right side of Eq. 1 into the right side of Eq. 30.6, I have $2\pi RB = \mu_0 I$ and so $B = \mu_0 I/(2\pi R)$, which is the same value I got using Eq. 30.1. ✔

❹ **EVALUATE RESULT** It's reassuring to see that the magnetic field magnitude at P does not depend on the choice of surface spanning the integration path. I can easily modify the argument above to show that any other surface that intercepts the left plate gives the same result. For example, a surface that intercepts one-quarter of the plate, as in **Figure 30.27**, intercepts one-quarter of the electric flux, and so the displacement current is only $I/4$. This surface intercepts the current twice: Where the surface intersects the wire, the current I crosses the surface from left to right, and where the surface intersects the plate, one-quarter of the current crosses the surface in the other direction, for a total contribution of $\frac{3}{4}I$. Again the sum of the ordinary and displacement currents intercepting the surface is I.

Figure 30.26

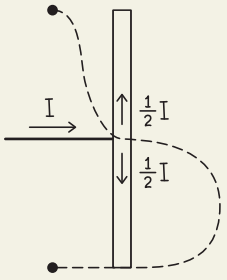
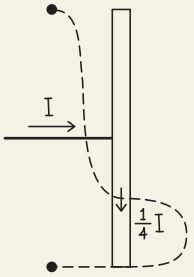


Figure 30.27



EXAMPLE 30.5 Magnetic field in a capacitor

A parallel-plate capacitor has circular plates of radius $R = 0.10$ m and a plate separation distance $d = 0.10$ mm. While a current charges the capacitor, the magnitude of the potential difference between the plates increases by 10 V/μs. What is the magnitude of the magnetic field between the plates at a distance R from the horizontal axis of the capacitor?

❶ **GETTING STARTED** As the capacitor is charging, there is a changing electric flux between the plates, so the electric field between the plates is changing. This changing electric field is accompanied by a magnetic field.

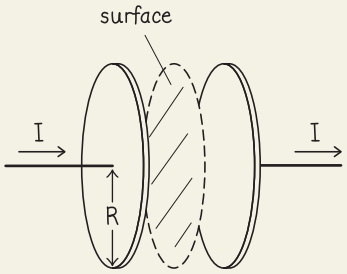
❷ **DEVISE PLAN** Equation 30.6 relates the magnetic field to a changing electric flux. To work out the left side of Eq. 30.6, I chose a circular integration path centered on the horizontal axis of the capacitor so that I can exploit the circular symmetry of the problem. To work out the right side of Eq. 30.6, I need to determine the rate of change of the electric flux through an appropriate surface spanning the integration

(Continued)

QUANTITATIVE TOOLS

path. I choose the simplest possible surface: a flat surface parallel to the plates (**Figure 30.28**). To calculate the change in electric flux intercepted by this surface, I need to determine the magnitude of the (uniform) electric field between the plates. In Example 26.2 I determined that the magnitude of the electric field between the plates is related to the plate separation distance d and the magnitude of the potential difference between the plates: $V_{cap} = Ed$.

Figure 30.28



3 EXECUTE PLAN Because the electric field between the plates is uniform and perpendicular to the plates, the electric flux Φ_E through the surface I chose is $\Phi_E = EA = E\pi R^2$, where $A = \pi R^2$ is also the area of the capacitor plates. The time rate of change of the electric flux is then given by

$$\frac{d\Phi_E}{dt} = \pi R^2 \frac{dE}{dt}.$$

Substituting this result in Eq. 30.6 and setting the current term equal to zero because no current is intercepted by the surface I've chosen, I get

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 \epsilon_0 \pi R^2 \frac{dE}{dt}.$$

Around my integration path, the magnitude of the magnetic field is constant, and the left side of this expression simplifies to $2\pi RB$. Solving for B, I obtain

$$B = \frac{\mu_0 \epsilon_0 R}{2} \frac{dE}{dt}. \qquad (1)$$

Because $V_{cap} = Ed$, I can write $E = V_{cap}/d$ and so

$$B = \frac{\mu_0 \epsilon_0 R}{2d} \frac{dV_{cap}}{dt} \qquad (2)$$

$$B = (4\pi \times 10^{-7}\ \text{T}\cdot\text{m/A})\,(8.85 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2) \times \frac{0.10\ \text{m}}{2(0.10 \times 10^{-3}\ \text{m})} \frac{10\ \text{V}}{1.0 \times 10^{-6}\ \text{s}}$$
$$= 5.6 \times 10^{-8}\ \text{T},$$

where I have used the Eq. 25.16 definition of the volt, $1\ \text{V} \equiv 1\ \text{J/C} \equiv 1\ \text{N}\cdot\text{m/C}$, and Eq. 27.3, for the ampere, $1\ \text{C} = 1\ \text{A}\cdot\text{s}$, to simplify the units. ✔

4 EVALUATE RESULT The magnetic field magnitude I obtain is very small, in spite of the substantial rate at which the potential difference between the plates increases. I have no way of knowing whether my numerical result is reasonable or not, but what I can do to evaluate the result is use another method to obtain an expression for B. Because my flat surface intercepts all the electric flux, I know that the magnitude of the magnetic field should be the same at all positions a distance R from the current-carrying wire. I can obtain the current by solving Eq. 26.1, $q/V_{cap} = C$, for the charge q on the capacitor plate and then using the definition of current, $I \equiv dq/dt$:

$$I \equiv \frac{dq}{dt} = C\frac{dV_{cap}}{dt}.$$

Substituting the capacitance of a parallel-plate capacitor $C = \epsilon_0 A/d = \epsilon_0 \pi R^2/d$ (see Example 26.2) into this expression, and then substituting the result into the expression for the magnetic field around a current-carrying wire from Example 28.3, $B = \mu_0 I/(2\pi r)$ (setting $r = R$ for the distance to the wire), I get

$$B = \frac{\mu_0 I}{2\pi R} = \frac{\mu_0 \epsilon_0 R}{2d} \frac{dV_{cap}}{dt},$$

the same result I obtained in Eq. 2, as I expect.

Note that in Example 30.5 the rate of change of the electric field is very large (about $10^{11}\ \text{V}/(\text{m}\cdot\text{s})$), but the accompanying magnetic field is small. This is not the case for electric fields that accompany changing magnetic fields, as substantial emfs can be induced by the motion of ordinary magnets.

30.11 Consider again the parallel-plate capacitor of Figure 30.23. For circular plates of radius R, calculate the magnitude of the magnetic field a distance $r < R$ from the horizontal axis of the capacitor (*a*) between the plates and (*b*) a short distance to the right of the right plate.

QUANTITATIVE TOOLS

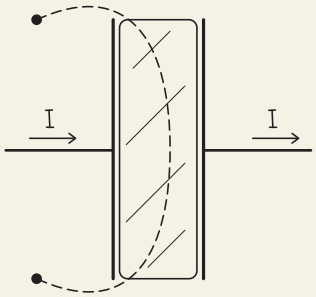
EXAMPLE 30.6 Displacement current in the presence of a dielectric

Suppose a slab of dielectric with dielectric constant κ is inserted between the plates of the capacitor in Figure 30.23 and the capacitor is charged with a current I, as considered in Example 30.1. How does Eq. 30.6 have to be modified to account for the dielectric?

❶ **GETTING STARTED** For a given amount of charge on the capacitor plates, the presence of a dielectric decreases the magnitude of the electric field between them. As I concluded in Example 30.1, however, the magnetic field surrounding the wires that lead to the capacitor wires is determined only by the current I through the wires and therefore cannot be affected by the insertion of the dielectric.

❷ **DEVISE PLAN** Given that the magnetic field surrounding the wires cannot be affected by the presence of the dielectric, the displacement current intercepted by a surface spanning a circular path around the wire and passing between the capacitor plates (**Figure 30.29**) should be equal to I, regardless of the presence of the dielectric. By setting the displacement current through the surface in Figure 30.29 equal to I, I can determine how the right side of Eq. 30.6 needs to be modified to account for the presence of the dielectric.

Figure 30.29



❸ **EXECUTE PLAN** As the capacitor charges, the presence of the dielectric reduces the magnitude of the electric field by a factor $1/\kappa$ (see Eq. 26.15): $E = E_{\text{free}}/\kappa$. As a result, the rate of change of the electric field dE/dt and the rate of change in the electric flux intercepted by the surface $d\Phi_E/dt$ are also reduced by a factor $1/\kappa$. To compensate for this reduction, I need to multiply $d\Phi_E/dt$ by κ in order to make the right side of Eq. 30.4 equal to I again. Therefore Eq. 30.6 becomes

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I + \mu_0 \epsilon_0 \kappa \frac{d\Phi_E}{dt}. \checkmark$$

❹ **EVALUATE RESULT** My modification to Eq. 30.6 is identical to the modification we made to make Gauss's law work in dielectrics (Eq. 26.25): In both cases the term containing the electric field or electric flux includes the factor κ.

30.5 Maxwell's equations

With Maxwell's addition of the displacement current $\epsilon_0 d\Phi_E/dt$ to Ampère's law, $\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I$, we now have a complete mathematical description of electric and magnetic phenomena and the relationship between the two. Let us summarize this description in the absence of a dielectric.

Electric and magnetic fields are *defined* by Eq. 27.20, which gives the force exerted on a charged particle moving in an electric field and a magnetic field:

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}). \tag{30.9}$$

Charged particles are the source of electrostatic fields, and electrostatic field lines always begin or end on charged particles. We can calculate electric fields from each individual charged particle if we wish, but when dealing with a distribution of charge that exhibits a certain symmetry (Section 24.4), it is most convenient to use Gauss's law (Eq. 24.8) to work out the electric field. Gauss's law tells us that the electric flux Φ_E through a closed surface (a Gaussian surface) is proportional to the charge enclosed by the surface:

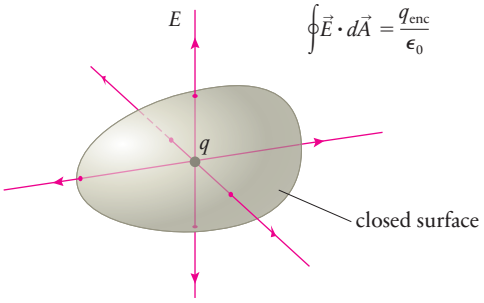
$$\Phi_E \equiv \oint \vec{E} \cdot d\vec{A} = \frac{q_{\text{enc}}}{\epsilon_0}. \tag{30.10}$$

Figure 30.30*a* shows the electric field created by a charged particle, along with a Gaussian surface enclosing that particle.

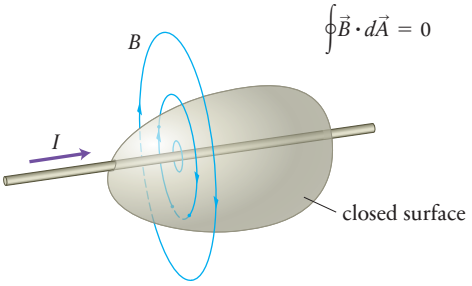
Magnetic fields are generated by moving charged particles, commonly in the form of currents. Unlike electric field lines, magnetic field lines always form loops. There are no isolated magnetic poles, only magnetic

Figure 30.30 Graphical representation of the physics behind Maxwell's equations, together with their mathematical expressions. (*a*) Electric field surrounding a charged particle and a Gaussian surface enclosing that particle. Gaussian surfaces can be used to relate the electric field to the enclosed charge. (*b*) Magnetic field surrounding a current-carrying wire and a closed surface intercepted by the wire; the integral of the magnetic field over a closed surface is always zero. (*c*) Electrostatic field and two closed paths through that field; the path integral of the electric field around either path must be zero. (*d*) Steady magnetic field surrounding a current and two closed paths through that field; the path integral of the magnetic field is proportional to the encircled current. (*e*) Changing magnetic field and two closed paths through it; an electric field accompanies the changing magnetic field; the path integral of this electric field around either path is nonzero. (*f*) Changing electric field and two closed paths through it; a magnetic field accompanies the changing electric field; the path integral of this magnetic field around either path is nonzero. Assume that no current is enclosed by either path.

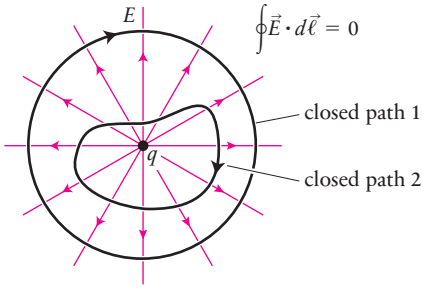
(*a*) Surface integral of electric field (Gauss's law)



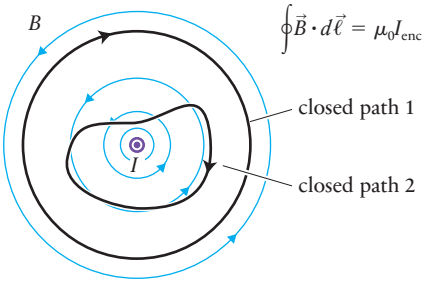
(*b*) Surface integral of magnetic field (Gauss's law for magnetism)



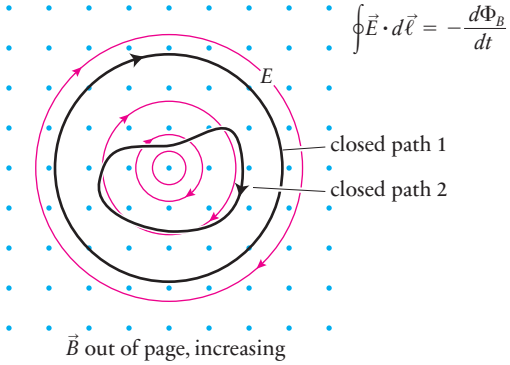
(*c*) Line integral of constant electric field



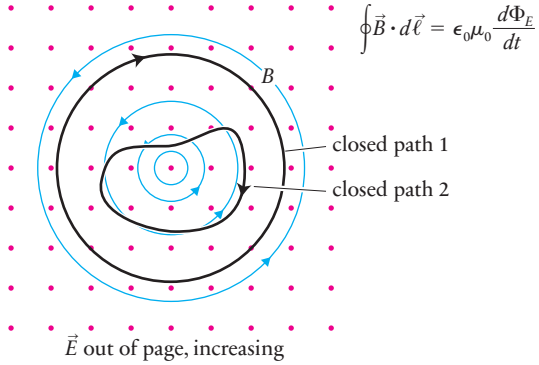
(*d*) Line integral of constant magnetic field (Ampère's law)



(*e*) Line integral of changing electric field (Faraday's law)



(*f*) Line integral of changing magnetic field (Maxwell's displacement current)



QUANTITATIVE TOOLS

dipoles. Consequently, as we showed in Chapter 27, the magnetic flux through any closed surface is always zero (Eq. 27.11):

$$\Phi_B \equiv \oint \vec{B} \cdot d\vec{A} = 0. \tag{30.11}$$

Figure 30.30*b* shows the magnetic field surrounding a current-carrying wire, along with a closed surface intercepted by the wire.

For electrostatic fields, we showed in Chapter 25 that the path integral of the electric field around a closed path is zero, which means that when a charged object is moved around a closed path in an electrostatic field, the work done on it is zero (Eq. 25.32). This situation is represented in Figure 30.30*c*. However, the electric field accompanying a changing magnetic field does work on charged particles even when those particles travel around closed paths:

$$\oint \vec{E} \cdot d\vec{\ell} = -\frac{d\Phi_B}{dt}. \tag{30.12}$$

Figure 30.30*e* shows the electric field lines associated with a changing magnetic field. For such an electric field, no potential can be defined because the path integral of the electric field depends on the path chosen.

Finally, Ampère's law gives the line integral of the magnetic field produced by a current (Figure 30.30*d*). The magnetic field that accompanies a changing electric field forms loops around the direction of the change in electric field, as shown in Figure 30.30*f*. Combining these two contributions to the line integral of the magnetic field gives us Maxwell's generalization of Ampère's law, which is our Eq. 30.6, repeated here:

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I_{\text{int}} + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}. \tag{30.13}$$

Equations 30.10–30.13 are referred to as **Maxwell's equations** because Maxwell not only added the displacement current term to Ampère's law but also recognized the coherence and completeness of this set of equations. Together with conservation of charge, these four equations give a complete description of electromagnetic phenomena. In the presence of matter, these equations have to be modified to account for the effects of matter on electric and magnetic fields (see, for example, Example 30.6).

Maxwell's equations were developed from and subsequently verified by a vast body of experimental evidence. Equation 30.10 (Gauss's law) comes from the measured inverse-square dependence of the electric force on separation distance and the finding that, in the steady state, the interior of a hollow charged conductor carries no surplus charge. Equation 30.11 (Gauss's law for magnetism) states that isolated magnetic monopoles do not exist, and none have been detected to date, in spite of very sensitive experiments conducted to search for them. Equation 30.12, a quantitative statement of Faraday's law, comes from extensive experiments by Faraday and others on electromagnetic induction, and Eq. 30.13, Maxwell's generalization of Ampère's law, comes from measurements of the magnetic force between current-carrying wires and the observed properties of electromagnetic waves.



30.12 Suppose that isolated magnetic monopoles carrying a "magnetic charge" m did exist, and that the interaction between these monopoles depended on $1/r^2$, where r is the distance between two monopoles. How would you modify Maxwell's equations to account for these monopoles? Ignore any physical constants that may need to be added.

QUANTITATIVE TOOLS

EXAMPLE 30.7 Maxwell's equations in free space

What is the form of Maxwell's equations in a region of space that does not contain any charged particles?

❶ **GETTING STARTED** If there are no charged particles, there can be no accumulation of charge and no currents, which means $q_{\text{enc}} = 0$ and $I = 0$.

❷ **DEVISE PLAN** All I need to do is set q_{enc} and I equal to zero in Eqs. 30.10–30.13.

❸ **EXECUTE PLAN** Setting $q_{\text{enc}} = 0$ in Eq. 30.10 and $I = 0$ in Eq. 30.13, I obtain the following form of Maxwell's equations:

$$\oint \vec{E} \cdot d\vec{A} = 0 \tag{1}$$

$$\oint \vec{B} \cdot d\vec{A} = 0 \tag{2}$$

$$\oint \vec{E} \cdot d\vec{\ell} = -\frac{d\Phi_B}{dt} \tag{3}$$

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}. \checkmark \tag{4}$$

❹ **EVALUATE RESULT** Maxwell's equations simplify greatly in the absence of charged particles (the only asymmetry is the sign difference between Eqs. 3 and 4, which comes from Lenz's law). Equations 1 and 2 state that both the electric and magnetic fluxes through a closed surface are zero in the absence of charged particles. Consequently, both electric and magnetic field lines must form loops. Equations 3 and 4 state that electric field line loops accompany changes in magnetic flux and magnetic field line loops accompany changes in electric flux.



30.13 As you saw in Section 30.3, the magnetic and electric fields in an electromagnetic wave are perpendicular to each other. How do Maxwell's equations in free space (Eqs. 1–4 of Example 30.7) express that perpendicular relationship?

30.6 Electromagnetic waves

From Maxwell's equations, we can derive the fundamental properties of electromagnetic waves. To begin, let's consider an electromagnetic wave pulse that arises from the sudden acceleration of a charged particle, as we discussed in the first part of this chapter (**Figure 30.31**). The magnitude of the electric field in the kinked part of the pulse in Figure 30.31 is essentially uniform and much greater than it is anywhere else. We shall consider the propagation of this wave pulse through a region of space containing no matter and no charged particles, so we can use the form of the Maxwell equations derived in Example 30.7. At great distances from the particle, only the transverse pulse is significant and we can ignore any other contributions to the electric field. This wave pulse is essentially a slab-like region of space that extends infinitely in the x and y directions and has a finite thickness in the z direction. Inside the slab, the electric field is uniform and has magnitude E; outside the slab, $E = 0$. We let the wave pulse move along the z axis (**Figure 30.32**) and denote its speed by c_0 (the subscript 0

Figure 30.31 Electric field pattern of an accelerated charged particle. The kinks in the electric field pattern correspond to a transverse electric field pulse propagating away from the particle at speed c_0.

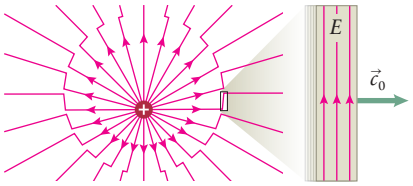
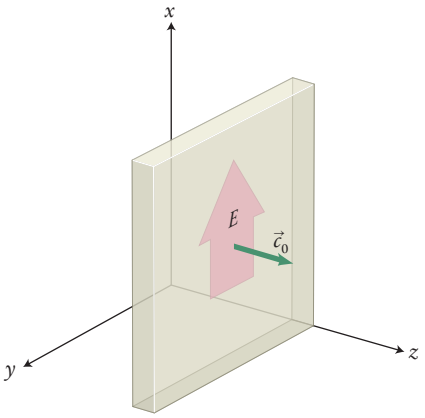


Figure 30.32 Perspective view of a planar electromagnetic wave pulse moving in the $+z$ direction. The electric field points in the $+x$ direction and has the same magnitude throughout an infinite plane parallel to the xy plane.



indicates that this speed is in vacuum). The magnitude of the electric field depends only on z, not on x and y. The wave pulse is an example of a **planar electromagnetic wave** because of the constant magnitude of the electric field in a plane normal to the direction of propagation.

What magnetic field pattern is associated with the electric field in the planar electromagnetic wave pulse in Figure 30.32? Viewing Figure 30.32 from the side (**Figure 30.33**), we see that the electric field is zero in front and in back of the pulse, nonzero and uniform inside the pulse, and changing at the front and back surfaces of the pulse. At the front surface, the electric field increases in the upward direction, corresponding to an upward displacement current I_{disp} (**Figure 30.34**). At the back surface of the pulse, the displacement current points down.

In Checkpoint 28.13*b*, you determined the magnetic field of two infinite planar sheets of oppositely directed current. The electric field in the planar electromagnetic wave pulse gives a similar arrangement of oppositely directed displacement currents. The magnetic field associated with this current distribution is uniform and points in the $+y$ direction (**Figure 30.35**). We now see that the planar electromagnetic wave pulse consists of uniform electric and magnetic fields that are perpendicular to each other and to the direction of propagation of the pulse, as we already concluded in Section 30.3. In fact, for a planar electromagnetic wave pulse, there is a right-hand relationship among the directions of $\vec{E}$, $\vec{B}$, and $\vec{c}_0$. If you curl the fingers of your right hand from the direction of $\vec{E}$ to the direction of $\vec{B}$, in Figure 30.35, your thumb points in the direction of propagation of the pulse. (This means that the vector product $\vec{E} \times \vec{B}$ yields a vector pointing in the direction of propagation of the electromagnetic wave pulse.)

To calculate the magnitude of the magnetic field in the planar electromagnetic wave pulse, we can use the version of Eq. 30.13 valid in a region of space that does not contain any charged particles (see Example 30.7):

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}. \tag{30.14}$$

Let's begin by evaluating the left side of this equation. To exploit the fact that the magnetic field points in the $+y$ direction in the pulse, we choose the Ampèrian path in Figure 30.35. This rectangular path lies in the yz plane and has width ℓ in the y direction; side ad is inside the pulse and side fg is far off to the right in the positive z direction. We let the

Figure 30.33 Side view of the planar electromagnetic wave pulse of Figure 30.32. The electric field inside the pulse is uniform except at the front and back surfaces, where it drops rapidly to zero.

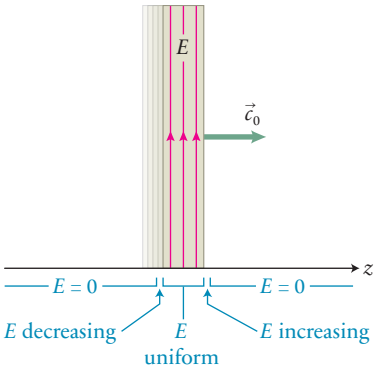
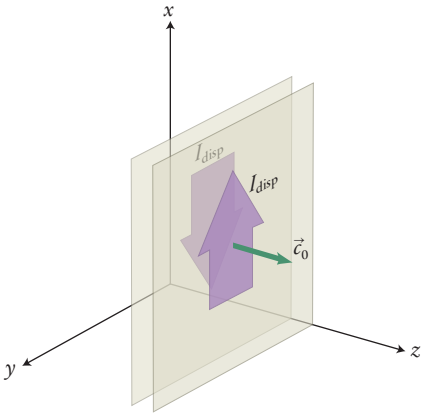


Figure 30.34 Displacement currents corresponding to the upwardly increasing electric field at the front surface and upwardly decreasing electric field at the back surface of the planar electromagnetic wave pulse of Figure 30.32.



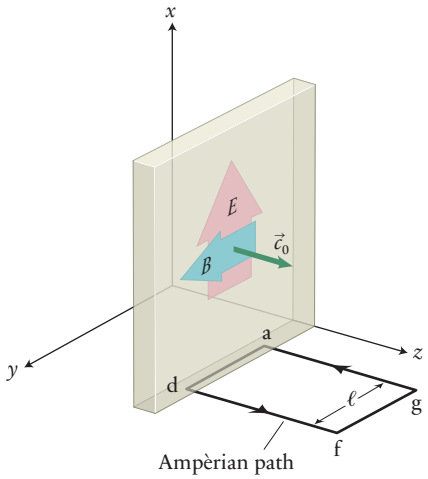
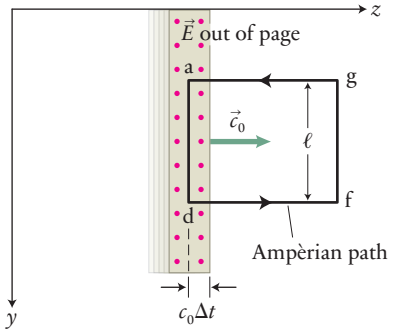


Figure 30.35 Magnetic field associated with the planar electromagnetic wave pulse of Figure 30.32. The Ampèrian path in the yz plane can be used to calculate the magnitude of the magnetic field.

direction of the path be such that it coincides with the direction of the magnetic field, so that $\vec{B} \cdot d\vec{\ell} = B\, d\ell$. Only side ad of the rectangular path contributes to the line integral; the magnetic field is zero around side fg, and the two long sides are perpendicular to the magnetic field. Thus, the left side of Eq. 30.14 becomes

$$\oint \vec{B} \cdot d\vec{\ell} = B\ell. \tag{30.15}$$

Figure 30.36 Top view of moving planar electromagnetic wave pulse of Figures 30.32–30.35, showing motion of the pulse through the Ampèrian path.



The electric flux through the path is given by $\Phi_E = \vec{E} \cdot \vec{A} = EA$, where $\vec{A}$ is a surface area vector pointing in the $+x$ direction, as dictated by the choice of direction of the integration path (Appendix B). To determine the rate of change of the electric flux through the path, note that the planar electromagnetic wave pulse is moving to the right with speed c_0. Before the front surface of the pulse reaches side ad of the Ampèrian path, the electric flux Φ_E through the path is zero. In a time interval Δt after the front surface of the pulse reaches side ad, the pulse travels a distance $c_0\Delta t$ into the rectangular path (**Figure 30.36**), and so at this instant the area over which the electric field is nonzero is $A = \ell c_0 \Delta t$. The electric flux through the path is then $\Phi_E = E\ell c_0 \Delta t$. The change in electric flux through the path during the interval Δt is thus

$$\Delta\Phi_E = E\ell c_0 \Delta t - 0, \tag{30.16}$$

and the rate of change in electric flux is

$$\frac{\Delta\Phi_E}{\Delta t} = E\ell c_0. \tag{30.17}$$

Substituting Eqs. 30.15 and 30.17 into Eq. 30.14 yields

$$B\ell = \mu_0 \epsilon_0\, E\ell c_0 \tag{30.18}$$

or, solving for B,

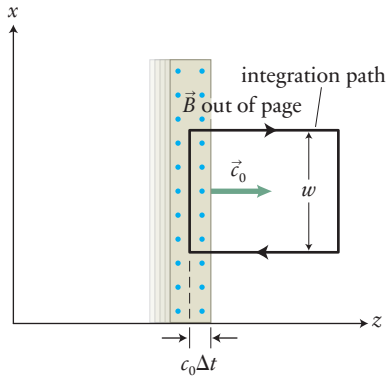
$$B = \mu_0 \epsilon_0\, E c_0. \tag{30.19}$$

We have thus obtained a relationship between the magnitudes of the magnetic and electric fields in the planar electromagnetic wave pulse.

We can now use Faraday's law (Eq. 30.12) to obtain an additional relationship between these two transverse fields:

$$\oint \vec{E} \cdot d\vec{\ell} = -\frac{d\Phi_B}{dt}. \tag{30.20}$$

Figure 30.37 Top view of the planar electromagnetic wave pulse of Figures 30.32–30.35, showing motion of the pulse through an integration path lying in the xz plane.



Let's begin by evaluating the left side of this equation. To exploit the fact that the electric field points in the $+x$ direction in the pulse, we choose the rectangular integration path in **Figure 30.37**. We let the direction of the path be such that it coincides with the direction of the electric field in the pulse, so that $\vec{E} \cdot d\vec{\ell} = E\, d\ell$. As in our derivation of the magnetic field, the only contribution to the line integral of the electric field around the rectangular path comes from the left side of the path:

$$\oint \vec{E} \cdot d\vec{\ell} = Ew. \tag{30.21}$$

To evaluate the right side of Eq. 30.20, we note that the geometry of this situation is the same as in our treatment of the electric pulse, except that now $\Phi_B = \vec{B} \cdot \vec{A} = -BA$, where $\vec{A}$ is a surface area vector pointing in

QUANTITATIVE TOOLS

the $-y$ direction, as dictated by the choice of direction of the integration path (see Appendix B). In analogy to Eqs. 30.16 and 30.17, the rate at which the magnetic flux through the path changes is then given by

$$\frac{\Delta\Phi_B}{\Delta t} = -Bwc_0. \tag{30.22}$$

Substituting Eqs. 30.21 and 30.22 into Eq. 30.20 gives us

$$Ew = Bwc_0 \tag{30.23}$$

or

$$B = \frac{E}{c_0}. \tag{30.24}$$

We now have two different relationships between the magnitudes of the electric and magnetic fields: Eq. 30.19 comes from Maxwell's generalization of Ampère's law, and Eq. 30.24 comes from Faraday's law. Setting the right sides of these two equations equal, we get

$$\frac{E}{c_0} = \mu_0\epsilon_0 Ec_0. \tag{30.25}$$

This result implies that the speed of the planar electromagnetic wave pulse is

$$c_0 = \frac{1}{\sqrt{\epsilon_0\mu_0}}. \tag{30.26}$$

Equation 30.26 tells us something surprising: The speed of the planar electromagnetic wave pulse in vacuum is determined by two fundamental constants, ϵ_0 and μ_0. The first, ϵ_0, is introduced in Coulomb's law (see Eqs. 22.1 and 24.7). The second, μ_0, (see Eq. 28.1), is introduced in Ampère's law. In 1862, when Maxwell first worked out the relationship expressed in Eq. 30.26, no one knew that light and electromagnetic waves were related. To evaluate c_0 in Eq. 30.26, Maxwell used the results of experiments made with electric circuits and obtained a value of $c_0 = 3 \times 10^8$ m/s, in excellent agreement with values obtained for the speed of light in vacuum. This agreement led Maxwell to the remarkable conclusion that light is an electromagnetic wave.

Nowadays, the speed of light is set to be exactly 299,792,458 m/s (see Section 1.3) to define the meter. The value of μ_0 is set to be exactly $4\pi \times 10^{-7}$ T·m/A (see Section. 28.5). The value of $\epsilon_0 = 1/(c_0^2\mu_0)$ as given in Eq. 24.7 is therefore also fixed.

Most electromagnetic waves have a more complex shape than the planar electromagnetic wave pulse we have used to arrive at Eq. 30.26. Through the superposition principle (see Section 16.3), however, we can superpose any number of planar electromagnetic wave pulses to obtain whatever planar electromagnetic wave shape interests us. The central property of these electromagnetic waves does not depend on shape: The electromagnetic wave pulse consists of electric and magnetic fields that are perpendicular to each other and to the direction of propagation of the pulse. The ratio of the magnitudes of the electric and magnetic fields is always given by Eq. 30.24. The field vectors $\vec{E}$ and $\vec{B}$ are always perpendicular, and they always travel at speed c_0 in a direction given by the vector product $\vec{E} \times \vec{B}$.

Mathematically, it is more convenient to build arbitrary wave shapes out of harmonic (sinusoidal) waves than out of rectangular wave pulses.

Figure 30.38 Perspective view of the electric field of a sinusoidal planar electromagnetic wave propagating in the z direction. The electric field vectors are embedded in rectangular slabs to emphasize that the electric field has the same magnitude everywhere in the plane of the slab. The magnitude of the electric field does vary from plane to plane along the z axis. The magnetic field (not shown) is uniform on planes parallel to the xy plane.

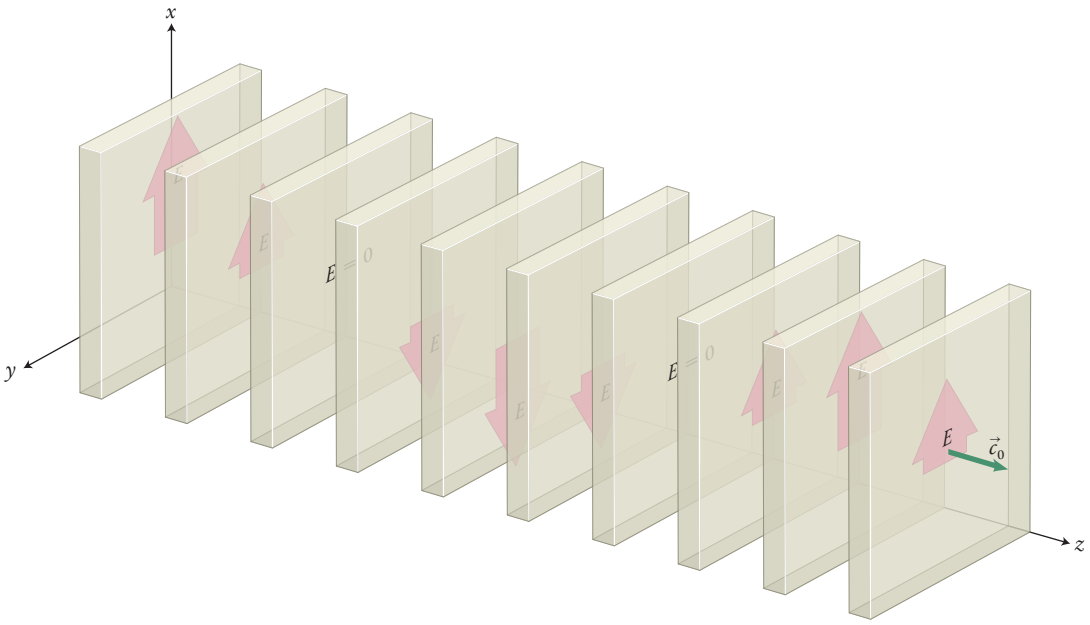


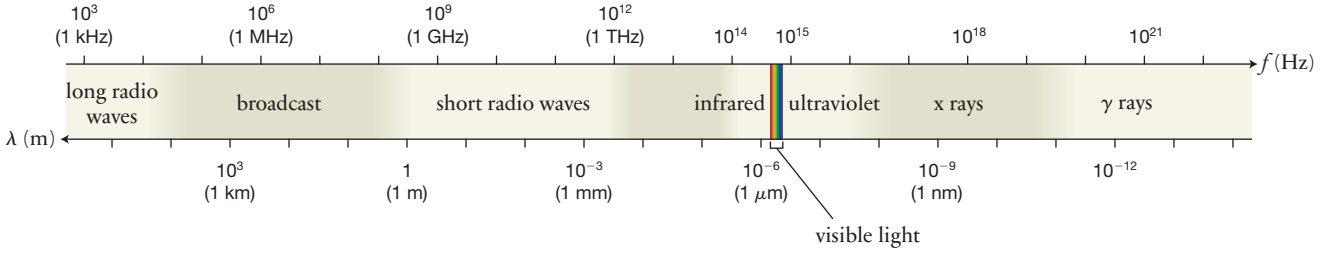
Figure 30.38 shows a planar electromagnetic wave for which the electric field varies sinusoidally in space. The field vectors for the electric field are shown embedded in rectangular slabs to emphasize that the electric field has the same magnitude everywhere throughout the plane of a slab, not just on the z axis. The magnitude of this electric field depends only on z, not on x and y.

As we saw in Chapter 16, harmonic waves are characterized by a propagation speed c, a frequency f, and a wavelength λ, and these quantities are related by $c = f\lambda$. The remarkable thing about electromagnetic waves is that waves of all frequencies travel at the same constant speed c_0 in vacuum. Consequently, in vacuum, frequency and wavelength are inversely proportional to one another over a vast range of values.

Figure 30.39 shows the classification of electromagnetic waves as a function of wavelength and frequency. Extending over a span of nearly 20 orders of magnitude, the figure shows electromagnetic waves ranging from radio waves to gamma rays. Only a very small part of this range corresponds to what we are familiar with as "light." Our eyes are most sensitive to wavelengths between 430 nm and 690 nm, though we can see light somewhat outside this wavelength range if the light is sufficiently intense. However, waves outside the visible range are governed by exactly the same physics as visible light.

QUANTITATIVE TOOLS

Figure 30.39 Classification of electromagnetic radiation as a function of frequency (top scale) and wavelength (bottom scale).



As we shall explore in more detail in Chapter 33, the frequency of an electromagnetic wave determines how the wave interacts with materials.

EXAMPLE 30.8 Speed of light in a dielectric

At what speed does an electromagnetic wave pulse propagate through a dielectric for which the dielectric constant is κ?

❶ **GETTING STARTED** In the presence of a dielectric, both Gauss's law and the displacement current are modified (Eq. 26.25 and Example 30.6, respectively). These changes affect the derivation I used to obtain the speed of electromagnetic waves in vacuum, Eq. 30.26.

❷ **DEVISE PLAN** The modification of the displacement current by the dielectric changes Eq. 30.14. To obtain an expression for the speed at which the electromagnetic wave pulse propagates, I carry the modified expression through the same logic I used to go from Eq. 30.14 to Eq. 30.26.

❸ **EXECUTE PLAN** Because there can be no conventional currents through a dielectric, I in Eq. 30.13 is still zero. Substituting the modified displacement current from Example 30.6, I get

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 \epsilon_0 \kappa \frac{d\Phi_E}{dt}.$$

Carrying this expression through the logic leading from Eq. 30.14 to Eq. 30.19, I obtain the magnitude of the magnetic field in terms of the electric field:

$$B = \mu_0 \epsilon_0 \kappa \, Ec, \tag{1}$$

where the κ comes from the modification of the displacement current and where I have written c for the speed in the dielectric rather than c_0, our symbol for the speed in vacuum.

Because Faraday's law is unaffected by the presence of the dielectric, the only modification required in Eq. 30.24 is replacing c_0 with c. Solving Eq. 30.24 for E and substituting that expression for E into Eq. 1, I get

$$B = \mu_0 \epsilon_0 \kappa \, Bc^2,$$

so the speed of an electromagnetic wave pulse moving through a dielectric is

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0 \kappa}} = \frac{c_0}{\sqrt{\kappa}}, \checkmark$$

where I have used Eq. 30.26, $c_0 = 1/\sqrt{\epsilon_0 \mu_0}$, to simplify.

❹ **EVALUATE RESULT** Because $\kappa > 1$ for most dielectrics, my result indicates that the speed of electromagnetic waves (including light) in a dielectric is smaller than their speed in vacuum. The dielectric constant is a measure of the reduction of the electric field (Eq. 26.15, $\kappa = E_{\text{free}}/E$). For a material that completely attenuates the electric field so that $E = 0$—in other words, for a conductor—κ is infinite and so my result yields $c = 0$. This means that an electromagnetic wave cannot propagate in such a material, a conclusion that agrees with the familiar observation that a conductor, such as a slab of metal, does not transmit visible light.



30.14 An electromagnetic wave with a wavelength of 600 nm in vacuum enters a dielectric for which $\kappa = 1.30$. What are the frequency and wavelength of the wave inside the dielectric?

30.7 Electromagnetic energy

Because electric and magnetic fields contain energy, energy is transported as an electromagnetic wave travels away from its source. Let us work out how much energy is transported by a planar electromagnetic wave. In Section 26.6, we calculated the energy density contained in an electric field in vacuum (Eq. 26.6):

$$u_E = \tfrac{1}{2} \epsilon_0 E^2. \tag{30.27}$$

Similarly, the energy density in a magnetic field in vacuum is given by Eq. 29.29:

$$u_B = \frac{1}{2} \frac{B^2}{\mu_0}. \tag{30.28}$$

The energy density in a combined electric and magnetic field is therefore

$$u = \tfrac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \frac{B^2}{\mu_0}. \tag{30.29}$$

Because the magnitudes of the electric and magnetic fields in an electromagnetic wave are related by Eq. 30.24, we can rewrite Eq. 30.29 in terms of just the magnitude of the electric field. Using Eqs. 30.24 and 30.26, we get

$$u = \tfrac{1}{2}\epsilon_0 E^2 + \tfrac{1}{2}\frac{E^2}{c_0^2\mu_0} = \tfrac{1}{2}\epsilon_0 E^2 + \tfrac{1}{2}\epsilon_0 E^2 = \epsilon_0 E^2. \tag{30.30}$$

Comparing Eqs. 30.29 and 30.30, we see that in vacuum the electric and magnetic fields each contribute half of the energy density—the electric and magnetic energy densities are equal.

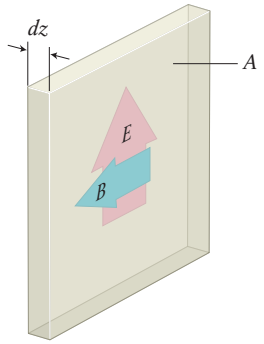
Alternatively, the energy density can be written in terms of only the magnitude of the magnetic field,

$$u = \frac{B^2}{\mu_0}, \tag{30.31}$$

or in terms of the magnitudes of both the electric and the magnetic fields,

$$u = \sqrt{\frac{\epsilon_0}{\mu_0}}EB. \tag{30.32}$$

Figure 30.40 Perspective view of a slice of thickness dz of an electromagnetic wave, taken normal to the direction of propagation.



Let us now calculate the rate at which energy flows through a certain area in an electromagnetic wave. Consider taking a slice of an electromagnetic wave normal to the direction of propagation (**Figure 30.40**). The slice has thickness dz and area A. The energy dU in this slice is the product of the energy density and the volume of the slice:

$$dU = uA\,dz. \tag{30.33}$$

From Section 17.5, you know that the intensity S of a wave is defined as the energy flow per unit time (the power) across a unit area perpendicular to the direction of wave propagation.* Using Eq. 30.33, we can express this relationship in the form

$$S = \frac{1}{A}\frac{dU}{dt}. \tag{30.34}$$

To determine the intensity of an electromagnetic wave, we substitute the expression for dU from Eq. 30.33 and recall that the wave travels at speed $dz/dt = c_0$. We can then rewrite Eq. 30.34 as

$$S = \frac{1}{A}uA\frac{dz}{dt} = uc_0 \tag{30.35}$$

or, after substituting Eqs. 30.32 and 30.26,

$$S = \frac{1}{\mu_0}EB. \tag{30.36}$$

*In Chapter 17, we used I for intensity. That is fine in a discussion of mechanical waves, but now that we must deal with current I so frequently, we switch to the symbol S for intensity.

As the electromagnetic wave travels, energy travels with it in the direction of wave propagation. As we saw in Section 30.6, the propagation direction is the same as that of the vector product $\vec{E} \times \vec{B}$. So with Eq. 30.36 we can define a vector that fully describes energy flow in the electromagnetic wave:

$$\vec{S} \equiv \frac{1}{\mu_0} \vec{E} \times \vec{B}. \tag{30.37}$$

This vector $\vec{S}$ is called the **Poynting vector,** after J. H. Poynting (1852–1914), the physicist who first defined this vector. The SI units of the Poynting vector are W/m^2 (Checkpoint 30.16). The Poynting vector represents the flow of energy in any combined electric and magnetic field, not just electromagnetic waves (Checkpoint 30.15), and the direction of $\vec{S}$ is the direction of energy flow.

When we describe electromagnetic waves, the magnitude S of the Poynting vector is called the *intensity of the electromagnetic wave* and, as noted in Eq. 30.34, is the instantaneous electromagnetic power (energy per unit time) crossing a unit area. For electromagnetic waves, this area is perpendicular to the direction of the vector product $\vec{E} \times \vec{B}$. To obtain the power P crossing a surface, we integrate the Poynting vector over the surface:

$$P = \int_{\text{surface}} \vec{S} \cdot d\vec{A}. \tag{30.38}$$

The results of this section give expressions for the *instantaneous* values of electromagnetic energy density (Eqs. 30.30–32), intensity (Eq. 30.36), and power (Eq. 30.38). Often, however, it is useful to consider average values over some time interval. Because the average of the square of a sine function is 1/2 (see Appendix B), the average value of $\sin^2 \omega t$ is 1/2. So, although sinusoidally oscillating electric and magnetic fields average to zero, their squares average to 1/2 of the square of their amplitudes: $(E^2)_{\text{av}} = \frac{1}{2} E_{\text{max}}^2$. The square roots of these average values are often referred to as the *root-mean-square values*, or rms values (see Eq. 19.21), and are represented by E_{rms} and B_{rms}:

$$E_{\text{rms}} \equiv \sqrt{(E^2)_{\text{av}}} \quad \text{and} \quad B_{\text{rms}} \equiv \sqrt{(B^2)_{\text{av}}}. \tag{30.39}$$

For sinusoidal electromagnetic waves, energy density, intensity, and power are proportional to the squares of a sine function, and so the average values of these quantities are related to the rms values of the fields in the same manner that the instantaneous values of energy density, intensity, and power are related to the instantaneous values of the fields. For example, Eq. 30.36 yields

$$S_{\text{av}} = \frac{1}{\mu_0} E_{\text{rms}} B_{\text{rms}} \quad \text{(sinusoidal electromagnetic wave)}. \tag{30.40}$$

EXAMPLE 30.9 Tanning fields

The average intensity S of the Sun's radiation at Earth's surface is approximately $1.0\ kW/m^2$. Assuming sinusoidal electromagnetic waves, what are the root-mean-square values of the electric and magnetic fields?

❶ **GETTING STARTED** The electric and magnetic fields in an electromagnetic wave each contribute half of the energy density of the wave (Eq. 30.29). That energy density is related to the wave's Poynting vector $\vec{S}$, whose magnitude S is the intensity of an electromagnetic wave.

❷ **DEVISE PLAN** I can rewrite Eq. 30.35 to obtain the average energy density u_{av} in terms of the average intensity S_{av}. I can then use Eqs. 30.30 and 30.31 to relate the average energy density to the rms values of the electric and magnetic fields.

❸ **EXECUTE PLAN** From Eq. 30.35, I have

$$u_{\text{av}} = \frac{S_{\text{av}}}{c_0} = \frac{1.0 \times 10^3\ \text{W/m}^2}{3.0 \times 10^8\ \text{m/s}} = 3.3 \times 10^{-6}\ \text{J/m}^3.$$

(Continued)

QUANTITATIVE TOOLS

To obtain the value of E_{rms}, I express the relationship between u_{av} and E_{rms} by analogy to Eq. 30.30:

$$u_{av} = \epsilon_0 E_{rms}^2$$

$$E_{rms} = \sqrt{\frac{u_{av}}{\epsilon_0}} = \sqrt{\frac{3.3 \times 10^{-6}\ \text{J/m}^3}{8.85 \times 10^{-12}\ \text{C}^2/\text{N} \cdot \text{m}^2}}$$

$$= 6.1 \times 10^2\ \text{V/m}, \checkmark$$

where I have used the fact that 1 N/C = 1 V/m.

For sinusoidal electromagnetic waves, the relationship between instantaneous values of the electric and magnetic field magnitudes holds for the rms values of the field magnitudes, so I can use Eq. 30.24 to determine the value of B_{rms}:

$$B_{rms} = \frac{E_{rms}}{c_0} = \frac{6.1 \times 10^2\ \text{N/C}}{3.0 \times 10^8\ \text{m/s}} = 2.0 \times 10^{-6}\ \text{T}. \checkmark$$

4 EVALUATE RESULT As I expect based on the large value of c_0, the value of B_{rms} in teslas is much smaller than that of E_{rms} in volts per meter. This electric field magnitude is also not large, particularly compared with the electric fields often obtained in capacitors. (Applying a potential difference of just 1 V to the capacitor in the next example would produce an electric field magnitude of 10,000 V/m!) However, this electric field is greater than the typical atmospheric electric field magnitude of roughly 100 V/m (which is due to charged particles in the atmosphere rather than to the Sun's radiation).



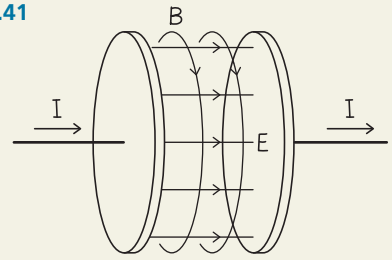
30.15 Consider supplying a constant current to a parallel-plate capacitor in which the plates are circular. While the capacitor is charging, what is the direction of the Poynting vector at points that lie on the cylindrical surface surrounding the space between the capacitor plates? What does this Poynting vector represent?

EXAMPLE 30.10 Capacitor power

A parallel-plate capacitor with circular plates of radius $R = 0.10$ m and separation distance $d = 0.10$ mm is charged by a constant current of 1.0 A. (*a*) What is the magnitude of the Poynting vector associated with the electric and magnetic fields at the edge of the space between the plates? (*b*) What is the rate at which electromagnetic energy is delivered to the cylindrical space between the plates?

1 GETTING STARTED While the capacitor is charging, the charge on the plates and consequently the electric field between them and the energy stored in the capacitor are changing with time. The changing electric field is accompanied by a magnetic field. The electric field between the plates is uniform, and the magnetic field lines form circular loops centered on the horizontal axis of the capacitor, in a direction given by the right-hand current rule. At each point in the space between the plates, the electric and magnetic fields are perpendicular to each other, as shown in my sketch in **Figure 30.41**.

Figure 30.41



2 DEVISE PLAN Because the electric and magnetic fields are perpendicular to each other, I can use Eq. 30.36 to calculate the magnitude of the Poynting vector. For this equation, I must know E and B. I calculated the magnitude of the electric field between two capacitor plates in Example 26.2, and from Checkpoint 30.11*a* I have an expression for the magnitude of the magnetic field between the plates.

To obtain the rate at which electromagnetic energy is delivered to the capacitor, I must integrate the Poynting vector over the cylindrical surface surrounding the space between the plates, as given in Eq. 30.38. From Checkpoint 30.15, I know that the Poynting vector points radially inward toward the capacitor axis, and I also know that at any given instant, the magnitudes of the electric and magnetic fields do not vary on the cylindrical surface surrounding the space between the plates. To integrate the Poynting vector over that surface, I therefore need to multiply the expression for the magnitude of the Poynting vector at any point on the surface by the area of the surface.

3 EXECUTE PLAN (*a*) From Example 26.2 I know that the electric field between the capacitor plates is $E = q/(\epsilon_0 A)$, where $A = \pi R^2$ is the area of each plate. Because a constant current is supplied to the capacitor while the capacitor is charging, $q = I\Delta t$ (Eq. 27.1, where Δt is the time interval that has elapsed since the capacitor began

charging), I can therefore rewrite my expression for E in the form

$$E = \frac{I\Delta t}{\epsilon_0 A} = \frac{I\Delta t}{\pi\epsilon_0 R^2}.$$

From the solution to Checkpoint 30.11*a* I have $B = \mu_0 Ir/2\pi R^2$, where r is the radial distance from the axis of the capacitor to the position where the magnetic field is measured and R is the radius of each plate. For this case, $r = R$, so

$$B = \frac{\mu_0 I}{2\pi R}.$$

To obtain the magnitude of the Poynting vector, I substitute the above values for E and B into Eq. 30.36:

$$S = \frac{EB}{\mu_0} = \frac{I^2\Delta t}{2\pi^2\epsilon_0 R^3}$$

$$= \frac{(1.0\text{ A})^2\Delta t}{2\pi^2[8.85\times10^{-12}\text{ C}^2/(\text{N}\cdot\text{m}^2)](0.10\text{ m})^3}$$

$$= [5.7\times10^{12}\text{ W}/(\text{m}^2\cdot\text{s})]\Delta t. \checkmark$$

(*b*) To calculate the rate at which electromagnetic energy is delivered to the cylindrical space between the capacitor plates, I multiply the expression for S by the area of the surface, which is the product of the circumference of the cylinder ($2\pi R$) and the height of the cylinder, which is the plate separation distance d:

$$P = 2\pi Rd\frac{I^2\Delta t}{2\pi^2\epsilon_0 R^3} = \frac{I^2 d\Delta t}{\pi\epsilon_0 R^2}$$

$$= \frac{(1.0\text{ A})^2(1.0\times10^{-4}\text{ m})\Delta t}{\pi[8.85\times10^{-12}\text{ C}^2/(\text{N}\cdot\text{m}^2)](0.10\text{ m})^2}$$

$$= (3.6\times10^8\text{ W/s})\Delta t. \checkmark$$

❹ **EVALUATE RESULT** I have no way of gauging the reasonableness of my numerical results, but I can check the reasonableness of my expression for the power (which incorporates my result for part *a*) by using an alternative method to obtain that expression. The rate at which energy is delivered to the capacitor is the power delivered to it by the current. I know from Eq. 26.2 that the electric potential energy of the capacitor changes by an amount $dU^E = V_{\text{cap}}\,dq$ during charging. To get the rate at which the amount of energy stored in the capacitor changes with time, I take the derivative

$$P \equiv \frac{dU_E}{dt} = V_{\text{cap}}\frac{dq}{dt} = V_{\text{cap}}I = \frac{qI}{C} = \frac{I^2\Delta t}{C}. \qquad (1)$$

For a parallel-plate capacitor, $C = \epsilon_0 A/d$, so with $A = \pi R^2$,

$$P = \frac{I^2 d\Delta t}{\epsilon_0\pi R^2},$$

which is the same result I obtained for part *b*.



30.16 Use Eq. 30.36 to show that the SI units of the Poynting vector are W/m^2.

Chapter Summary

Changing electric fields (Sections 30.1, 30.4)

Concepts A changing electric field is accompanied by a magnetic field, and a changing magnetic field is accompanied by an electric field.

Table 30.1 Properties of electric and magnetic fields

	Electric field	Magnetic field
associated with	charged particle	moving charged particle
	changing magnetic field	changing electric field
exerts force on	any charged particle	moving charged particle

Quantitative Tools The **displacement current** due to a changing electric flux through a surface is

$$I_{\text{disp}} \equiv \epsilon_0 \frac{d\Phi_E}{dt}. \tag{30.7}$$

The *Maxwell-Ampère law* is

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 (I_{\text{int}} + I_{\text{disp}}), \tag{30.8}$$

where I_{int} is the current intercepted by a surface spanning an Ampèrian path.

Maxwell's equations (Section 30.5)

Concepts **Maxwell's equations** provide a complete mathematical description of electric and magnetic phenomena and the relationship between the two.

Quantitative Tools **Maxwell's equations** are

$$\Phi_E \equiv \oint \vec{E} \cdot d\vec{A} = \frac{q_{\text{enc}}}{\epsilon_0} \tag{30.10}$$

$$\Phi_B \equiv \oint \vec{B} \cdot d\vec{A} = 0 \tag{30.11}$$

$$\oint \vec{E} \cdot d\vec{\ell} = -\frac{d\Phi_B}{dt} \tag{30.12}$$

$$\oint \vec{B} \cdot d\vec{\ell} = \mu_0 I_{\text{int}} + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}. \tag{30.13}$$

Electromagnetic waves (Sections 30.2, 30.3, 30.6, 30.7)

Concepts An **electromagnetic wave** is a propagating disturbance in a combined electric and magnetic field. The electric and magnetic fields in these waves are perpendicular to each other, and the direction of wave propagation is the direction of $\vec{E} \times \vec{B}$.

A **planar electromagnetic wave** is one in which the wavefronts are planes perpendicular to the direction of propagation. At each point on a planar wavefront, the instantaneous magnitude of the electric field has the same value, as does the instantaneous magnitude of the magnetic field.

Oscillating dipoles and antennas to which an alternating potential is applied emit electromagnetic waves.

The **polarization** of an electromagnetic wave is the orientation of the electric field as seen by an observer looking in the direction of propagation.

Quantitative Tools The speed of an electromagnetic wave in vacuum is

$$c_0 = \frac{1}{\sqrt{\epsilon_0 \mu_0}}. \tag{30.26}$$

The speed of an electromagnetic wave in a dielectric of dielectric constant κ is

$$c = \frac{c_0}{\sqrt{\kappa}}.$$

The **Poynting vector** $\vec{S}$ gives the energy transported by an electromagnetic wave per unit time across a unit area perpendicular to the direction of propagation of the wave:

$$\vec{S} \equiv \frac{1}{\mu_0} \vec{E} \times \vec{B}. \tag{30.37}$$

The power transported by an electromagnetic wave crossing a surface is

$$P = \int_{\text{surface}} \vec{S} \cdot d\vec{A}. \tag{30.38}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

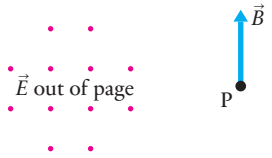
Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

Unless directed otherwise, use $c_0 = 3.0 \times 10^8$ m/s for the speed of electromagnetic waves traveling in vacuum.

30.1 Magnetic fields accompany changing electric fields

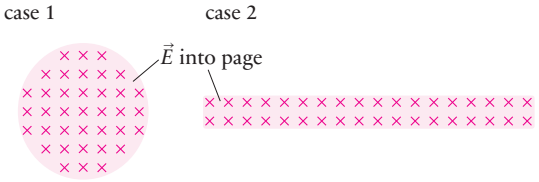
1. The magnitude of the electric field of Figure P30.1 is changing with time. As a result of this change, there is an upward-pointing magnetic field at position P. Is the electric field magnitude increasing or decreasing? •

Figure P30.1



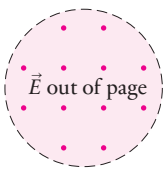
2. Figure P30.2 shows two electric fields, one in a region of circular cross-section and one in a long flat region. In both cases, the electric field decreases over time. What is the direction of the accompanying magnetic field in each case? •

Figure P30.2



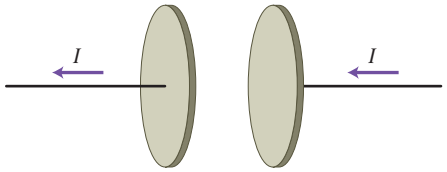
3. A parallel-plate capacitor is charged until it carries charge $+q$ on one plate and charge $-q$ on the other plate. The capacitor is then disconnected from the power supply and isolated. What is the direction of the magnetic field that surrounds the charged capacitor? •
4. The uniform electric field shown in Figure P30.4 points out of the page and is contained in a cylindrical volume of space. If the electric field is rapidly turned off, a magnetic field is produced. (*a*) What is the direction of this magnetic field? (*b*) Does this magnetic field exist only in the volume where the electric field exists, only outside that volume, or both inside and outside that volume? ••

Figure P30.4



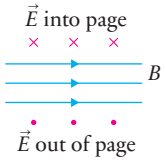
5. The capacitor of Figure P30.5 is being charged by a steady current I. What are the directions of (*a*) the electric field between the plates, (*b*) the change in the electric field, and (*c*) the magnetic field between the plates? ••

Figure P30.5



6. Figure P30.6 shows an increasing magnetic field and the electric field that accompanies it according to Faraday's law. If the magnetic field is increasing at an increasing rate, (*a*) is the electric field magnitude constant, increasing, or decreasing? (*b*) Does the behavior of the electric field produce an additional magnetic field? If so, what is the direction of the additional magnetic field? Does it reinforce or counteract the original magnetic field? ••

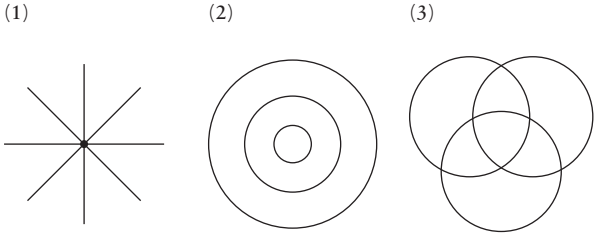
Figure P30.6



30.2 Fields of moving charged particles

7. Are the radiating electric field lines around a charged particle straight lines when the particle (*a*) is stationary, (*b*) moves at constant velocity, (*c*) accelerates? •
8. Which field line patterns in Figure P30.8 could represent (*a*) electric field lines and (*b*) magnetic field lines? •

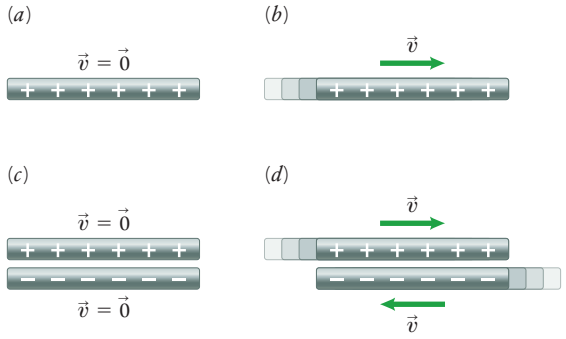
Figure P30.8



9. Would a charged particle moving at a high speed in a circle create kinks in its electric field? •
10. A positively charged particle moving to the right is stopped rapidly. Sketch the resultant radiated pulse. ••

11. Is any type of electric field produced when (*a*) a wire having zero surplus charge carries a constant current, (*b*) a wire having zero surplus charge carries a time-varying current, (*c*) a bar magnet moves translationally at constant velocity, (*d*) a bar magnet rotates about an axis that passes symmetrically through its poles? ●●
12. Is any magnetic field produced when (*a*) a uniformly charged sphere rotates at constant rotational speed about an axis that passes through its fixed center, (*b*) a charged particle moves translationally at constant velocity, (*c*) a charged particle that is moving translationally accelerates? ●●
13. Sketch the electric and magnetic field line patterns produced by the four situations shown in Figure P30.13: (*a*) a stationary charged rod carrying a linear charge density of $+\lambda$, (*b*) a charged rod carrying a linear charge density of $+\lambda$ that moves at constant velocity along the length of the rod, (*c*) a stationary charged rod carrying a linear charge density of $+\lambda$ that is infinitesimally close to a parallel stationary charged rod carrying a linear charge density of $-\lambda$, (*d*) a charged rod carrying a linear charge density of $+\lambda$ that moves at constant velocity along the length of the rod and is infinitesimally close to a parallel charged rod carrying a linear charge density of $-\lambda$ and moving in the opposite direction. ●●

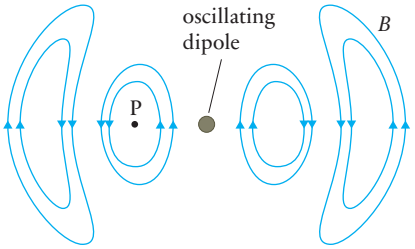
Figure P30.13



30.3 Oscillating dipoles and antennas

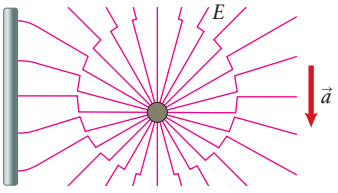
14. What antenna length should be used to create a dipole antenna that emits 100-MHz electromagnetic waves, a frequency typical of FM radio broadcasts? ●
15. An electromagnetic wave is traveling eastward. If at a given instant the magnetic field for one section of the wave points south, in which direction does the electric field for that section point at that instant? ●
16. An electric dipole with its center located at the origin of a Cartesian coordinate system oscillates along the x axis, creating an electromagnetic wave. At a position on the z axis far from the origin, (*a*) what is the polarization of the wave and (*b*) which axis are the magnetic field lines parallel to? ●●
17. Figure P30.17 shows a magnetic field line pattern radiating from an oscillating electric dipole. (*a*) What is the direction of the change, over the next infinitesimal time interval, in the magnetic field at P? (*b*) Use your answer to part *a* to determine the direction of the electric field loop around P. (*c*) Is the dipole oscillating right-left, up-down, or in-out of the page? ●●

Figure P30.17



18. A router used for wireless Internet access follows the 802.11g standard, which operates at 2.4 GHz. (*a*) What is the wavelength of the electromagnetic waves emitted by the router? (*b*) How tall would a dipole antenna need to be? Is this length comparable to the length of any router antennas you've seen? ●●
19. Why are radio transmission towers oriented vertically rather than laid horizontally on the ground? ●●
20. To detect an incoming planar electromagnetic wave, we observe a single charged particle that begins at rest and is free to move only in the xy plane. If we observe the particle oscillating along the y axis, with no motion along the x axis, what can we say about the direction of the electric and magnetic fields and the direction of wave propagation? ●●
21. In Figure P30.21, a charged particle has been suddenly accelerated downward, producing the electric field line pattern shown. When the kinks in the field lines reach the metal rod at the left, they produce a downward current in the rod. Is the particle that has been accelerated positively or negatively charged? ●●

Figure P30.21



22. As we saw in Checkpoint 30.9, a receiving dipole antenna has to be aligned with the oscillating electric field in order to produce a measurable potential difference in the antenna. What if we wanted to detect the potential difference caused by the changing magnetic field? (*a*) What antenna shape would we need? (*b*) How would this antenna need to be oriented relative to the wave in order to maximize its ability to detect the wave? (*c*) Use Faraday's law to explain whether this antenna is a better detector for low-frequency waves or for high-frequency waves. ●●●

30.4 Displacement current

23. A parallel-plate capacitor with circular plates has a steady charging current of 3.0 A. The wires into and out of the plates attach to the plate centers. If the radius of each plate is 5 cm and there is no dielectric between the plates, what is the magnetic field magnitude inside the plates at a distance of (*a*) 2 cm and (*b*) 10 cm from the symmetry axis of the capacitor? •

24. A parallel-plate capacitor of capacitance 10 μF is being charged by a voltage source $V = V_0 \cos(\omega t)$. Suppose that $V_0 = 240$ V and $\omega = 50\ \text{s}^{-1}$. (*a*) What is the maximum displacement current through a cross-sectional area inside the capacitor that is half the area of each plate? (*b*) How would this area have to be oriented relative to the electric field to get this maximum displacement current? •

25. Instead of a capacitor in a circuit, we can get the same effect by slicing a thick wire in two, making our cut perpendicular to the wire's long axis. If the wire diameter is 5 mm and we place the two parallel circular surfaces of the cut wire 0.02 mm apart, what is the magnetic field magnitude at a position that lies between the two surfaces and is 1.0 mm from the long central axis of the wire at the instant the current in the wire is 1.5 A? Assume the charge carriers that create the current are uniformly distributed over the cross-sectional area of the wire. ••

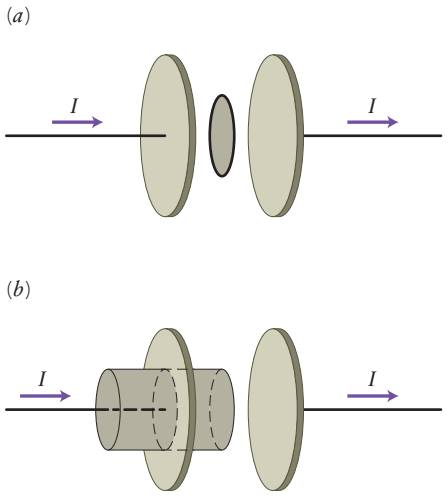
26. A parallel-plate capacitor has a steady charging current of 4.0 A. What is (*a*) the time rate of change of the electric flux between the plates and (*b*) the displacement current between the plates? (*c*) How do your answers to parts (*a*) and (*b*) change if there is a dielectric for which $\kappa = 2.2$ between the plates? Assume the area of the plates is much greater than the separation distance between them. ••

27. A parallel-plate capacitor with circular plates of radius r is charging with constant current I. At some instant, $B = 1.6 \times 10^{-8}$ T at a position 7.0 mm from the center of a long wire leading to the capacitor. (*a*) Calculate the current magnitude at that instant. (*b*) Draw the current in one plate to show why you get the same value of B no matter where you place the surface bounded by your Ampèrian path. Draw the lengths of your current arrows to scale to show how the current magnitude does or does not change at various positions in the plate. ••

28. For a capacitor of capacitance C, show that the displacement current between the capacitor's plates is given by $C(dV/dt)$, where V is the potential difference across the capacitor. ••

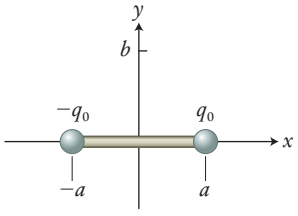
29. The circular plates of a parallel-plate capacitor have a radius of 30 mm. A steady 2.0-A current is charging the initially uncharged capacitor, and the surface charge on the plates is distributed uniformly. (*a*) Derive an expression for the magnitude of the electric field between the plates as a function of time. (*b*) In Figure P30.29*a*, the flat circular surface between the capacitor plates has a radius of 10 mm. What is the displacement current through this flat circular surface? (*c*) Consider the open surface shown in Figure P30.29*b*, a cylindrical shell open at the right end that passes through the left capacitor plate. If the radius of the flat circular end is 10 mm, what is the displacement current through the open shell? (*d*) Why is your answer to part *c* different from your answer to part *b*? Why are both these answers different from $I = 2.0$ A, the current through the capacitor? ••

Figure P30.29



30. A short wire runs along an x axis from $x = a$ to $x = -a$ (Figure P30.30). At $t = 0$, there is a small sphere carrying charge $+q_0$ at $x = a$ and a small sphere carrying charge $-q_0$ at $x = -a$. As the spheres discharge and a current is established in the wire, show that, for $t > 0$, the generalized form of Ampère's law (Eq. 30.13) and the Biot-Savart law (Eq. 28.12) give the same result for the magnitude of the magnetic field a distance b up the y axis. (Hint: Obtain the electric flux through a circular disk in the yz plane, centered at the origin. Because the electric field is not constant over the surface of the disk, you must divide the disk into a nested set of rings of various radii and integrate. A representative ring should have radius r and radial extent dr.) •••

Figure P30.30

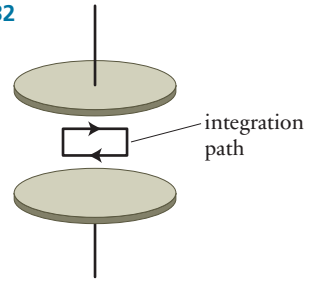


31. A parallel-plate capacitor has circular plates of radius $R = 15$ cm and plate separation distance $d = 1$ mm. While it is charging, the potential difference across the plates is given by $V(t) = V_0(1 - e^{-t/\tau})$ where $V_0 = 12$ V and $\tau = 4.0$ ms. What is the magnitude of the magnetic field between the plates at a distance $r = 5$ cm from the axis of the capacitor at $t = 2$ ms? •••

30.5 Maxwell's equations

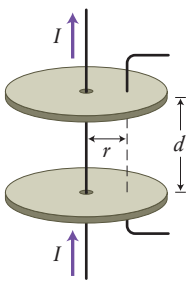
32. When you use Maxwell's equations to determine electric and magnetic fields, a lot depends on making the "right" choice for the integration path. What happens if you choose the "wrong" path? Do the equations still hold? Consider a parallel-plate capacitor with circular plates, but use a rectangular integration path that has two sides parallel to the plates and two sides perpendicular to them, as in Figure P30.32. What do you get? •

Figure P30.32



33. Using Eq. 30.11, show that the normal component of the magnetic field is continuous across any surface. ••
34. In free space, Maxwell's equations simplify greatly. The two equations involving surface integrals of the fields (Eqs. 30.10 and 30.11) are zero, and the two equations involving line integrals of the fields (Eqs. 30.12 and 30.13) have values proportional to the rate of flux change of the other field. Because the fluxes are surface integrals of the fields, why don't all four equations reduce to zero? ••
35. For a perfect conductor, $E = 0$ everywhere inside, and all the charge resides on the surface even when charge carriers are flowing. (*a*) Use Maxwell's equations to show that the magnetic field inside a conductor is constant. (*b*) Use Maxwell's equations to show that the magnetic flux through a perfect conducting loop is constant. (*c*) A superconductor is more than just a perfect conductor. It also has the property that $B = 0$ inside the superconductor. Use Maxwell's equations to show why the current in a superconductor is confined to the surface. •••
36. A physics teacher tries to build a device that illustrates Maxwell's generalization of Ampère's law. She drills a hole in the center of each plate of a parallel-plate capacitor and then runs a wire through the holes (Figure P30.36). There is no connection between the wire and the plates, and the potential difference across the plates can be controlled. Can the teacher adjust this potential difference so that there is no magnetic field at a distance r from the wire? In particular, what must be the rate of change of the potential difference if she wants no magnetic field at a distance $r = 20$ mm from the wire when it carries a steady current of 0.0200 A and the distance between the plates is $d = 3.0$ mm? •••

Figure P30.36



30.6 Electromagnetic waves

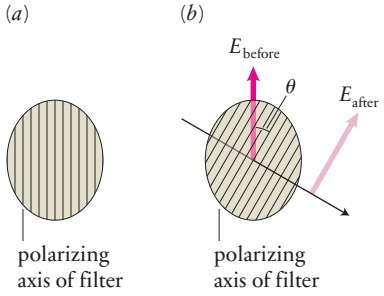
37. The frequency of AM radio channels is typically on the order of 10^5 Hz or 10^6 Hz. What is the wavelength of these waves? FM waves have wavelengths between 1.0 m and 10 m. What is the frequency of these waves? •
38. The speed of light drops to 2.26×10^8 m/s in water. What is the dielectric constant of water? •
39. The optical fibers used for telecommunications links have dielectric constant $\kappa = 1.6$. How long does a signal originating in California take to reach New York in a fiber-optic cable? •
40. A phone call you make from the United States to Germany is routed via a satellite that is in a geostationary orbit at an altitude of 36,000 km. How long is the time interval between the instant you say "Hello" to the instant your greeting arrives in Germany? Ignore the distance between the United States and Germany. •
41. The antenna of a WiFi access point has a length of about 80 mm. What does this say about the likely frequency of the wireless network? •
42. (*a*) At a distance from an emitting antenna where the electric field has a maximum magnitude of 15 V/m when the air is dry, what is the maximum magnetic field magnitude? (*b*) How does the value of E_{max} change when the air is very damp, increasing the dielectric constant of the air? (*c*) How does the value of B_{max} change in the damp air? Assume that the antenna's potential difference is unaffected by the damp air. ••
43. Using electromagnetic waves, you generally cannot resolve any structures that are smaller than the wavelength you are using. CD, DVD, and Blu-ray players use small "pits" that encode the 0s and 1s of the digital information stored on them—the smaller the pits, the more information you can store. A DVD can store 4.7 GB, and a Blu-ray can store 25 GB. If a DVD player typically uses lasers that emit red light in the wavelength range from 640 nm to 660 nm, what do you expect the wavelength and color of a Blu-ray laser to be? ••
44. The light produced by a sodium vapor lamp has a wavelength of 589.3 nm in vacuum. What is its wavelength after it enters a sheet of glass with dielectric constant $\kappa = 4.8$? ••
45. *3K cosmic background radiation* is energy left over from events that occurred when the universe was in a very early stage of development. Given that the amount of energy associated with any radiation is linearly proportional to the radiation frequency and that the energy associated with a 10^6-Hz radio wave is 6.6×10^{-28} J, what is the frequency of the cosmic background radiation? Assume that the energy of this radiation can be approximated as $k_B T$, where $T = 3.0$ K (thus the name given to this radiation). ••
46. If the electric field in an electromagnetic wave 5 cm from a radio-emitting antenna has a maximum magnitude of 3.0×10^5 V/m, what is the maximum magnetic field magnitude 1 km from the antenna? ••

30.7 Electromagnetic energy

47. An electromagnetic wave has an average Poynting vector magnitude of 4×10^{-7} W/m^2. What is the maximum value of the magnitude of the electric field? •

48. An electromagnetic wave has root-mean-square magnetic field magnitude $B_{rms} = 5\ \mu T$. What is the root-mean-square electric field magnitude and the average intensity of the wave? •
49. The power of a laser is much smaller than the power of an incandescent light bulb, but the laser light does not spread out very much. Explain why this tendency not to spread has such an effect over long distances. •
50. A 2.0-mW laser has a beam radius of 0.5 mm. What is the intensity of this beam? •
51. The intensity of the HERCULES laser, one of the world's most powerful, is $2.0 \times 10^{20}\ W/mm^2$. Granted, the beam pulse lasts for only 30 fs, but if we assume the beam is an electromagnetic pulse, what is its average energy density, and what energy can one pulse deliver to a 1.0-μm^2 target? ••
52. Radio signals typically have a very small intensity. Imagine that a vehicle receives an average signal of 20 $\mu W/m^2$. (*a*) What are the maximum magnitudes of the electric and magnetic fields? (*b*) If the tower emitting the radio waves is located 10 km away from the vehicle, how powerful is the emitting antenna? ••
53. For a constant current of 2 A, what time interval is required to deliver 0.5 MW of power to the space between the plates of a capacitor if the plates are circular and parallel, their diameter is 10 cm, and their separation distance is 1 mm? ••
54. Assume a 40-W incandescent light bulb radiates uniformly in all directions. At a distance of 1 m from the bulb, determine (*a*) the intensity of the electromagnetic waves, (*b*) the maximum electric field magnitude, and (*c*) the maximum magnetic field magnitude. ••
55. The emitting antenna of a 50-kW radio station radiates equally in all directions. What are the magnitudes E_{max} and B_{max} (*a*) 200 m from the antenna and (*b*) 20 km from the antenna? (*c*) For these two distances from the antenna, calculate the maximum potential difference caused between the ends of a receiving antenna that is 0.5 m long, assuming the antenna is aligned perfectly with the electric field of the radio wave. ••
56. A laser beam has a radius of 1 mm. How powerful does the laser have to be for the maximum magnitude of the magnetic field in the beam to be 4 μT? ••
57. Sunlight on Earth has an intensity of about $S = 1.0\ kW/m^2$. How much power can be harvested from a solar energy panel that is 2.0 m long and 1.6 m wide if the sunlight comes in at an angle of $\theta = 20°$ with respect to the surface normal vector of the panel and the panel is 20% efficient? ••
58. You want to use your microwave oven to heat 0.20 L of water from room temperature to boiling. The water is in a cup that has a radius of 30 mm, and you want the water to reach its boiling point in 3.00 min. What are the required (*a*) average power, (*b*) average intensity, and (*c*) root-mean-square electric field of the microwaves? Assume all the microwaves come in from the top and are completely absorbed by the water. ••
59. The human eye can detect light intensities as small as $S_{min} = 10^{-12}\ W/m^2$. If an incandescent 100-W light bulb is about 10% efficient, how far away could you see it under the simplest assumptions? Does your result make sense? ••
60. Similar to perceived loudness in hearing, the brightness the human eye and brain perceive for light is logarithmic rather than linear. For example, you do not perceive light that has an intensity of S as being twice as bright as light of intensity $S/2$. What happens to the perceived brightness of a light bulb if you double the distance from it? ••
61. A radio wave, with a wavelength of 300 m in vacuum, has an average intensity of 200 W/m^2. (*a*) What is the frequency of this electromagnetic wave (you may need to review material in Chapter 16)? (*b*) Fix a reference frame with the z axis along the direction of propagation of the wave. What is the Poynting vector of the wave as a function of position and time? (*c*) What is the time-averaged energy per unit time radiated through a 2.0-m^2 surface aligned with its normal parallel to the direction of propagation? (*d*) What are the electric and magnetic field vectors, as functions of position and time? •••
62. A cylindrical solenoid of radius R and height h consists of N windings. There is a current through the windings, and this current increases with time as $I = \alpha t$, where α is a constant. (*a*) Describe the directions of the electric field, the magnetic field, and the Poynting vector. (*b*) Determine the magnitude of the Poynting vector at a radial distance $r = R$ from the center of the long center axis of the solenoid. (*c*) Use the result of part *b* to derive an expression for the power input to the solenoid as a function of time. (*d*) Using the formula for the inductance of an ideal solenoid, calculate the power from the rate of change of the stored magnetic energy. Does this answer agree with your answer in part *c*? •••
63. A polarizing filter is a plastic sheet that allows only certain components of the electric field in an electromagnetic wave to pass through, and a wave that has passed through such a filter is said to be *polarized*. After an initially unpolarized wave has passed through the filter shown in Figure P30.63*a*, for instance, the wave's electric field contains only the vertical components of all the electric field vectors present before the wave passed through. Because the horizontal components have been blocked, the wave has lost half of its initial intensity. If this polarized wave then passes through the filter of Figure P30.63*b*, where the filter-polarizing direction is at an angle θ to the direction of the vertically polarized wave, the wave loses more of its intensity. If the polarized wave has an intensity $S_{before} = 200\ W/m^2$ before passing through the filter of Figure P30.63*b* and $\theta = 20.0°$, what is the intensity after the wave passes through? •••

Figure P30.63



Additional Problems

64. Liquid argon has a dielectric constant of 1.5. What is the speed of an electromagnetic wave pulse through this medium? •
65. What would be the length of a half-wave antenna for a radio station that emits its signal at 90.5 MHz? •
66. Full-body scanners at airports are sometimes referred to as *millimeter-wave scanners* and sometimes as *terahertz scanners*. Does this make sense? •
67. A home microwave oven typically uses 2.45-GHz microwaves, and a microwave oven in a restaurant kitchen is more likely to run at 915 MHz. Calculate the wavelength for each frequency and the length of the emitting antenna needed in each case. •
68. A sinusoidal electromagnetic wave has a maximum electric field magnitude of $E_{max} = 500$ N/C. What is the root-mean-square value of the electric field? •
69. By focusing an ultraviolet laser pulse, you can produce an apparently hovering plasma point in midair. If it takes an electric field magnitude of 1.0×10^6 N/C to ionize air, what must the intensity of the pulse be? ••
70. For a particular electromagnetic wave, B_{rms} is 0.30×10^{-6} T. For this wave, calculate (*a*) E_{rms} , (*b*) the average energy density, and (*c*) the intensity. ••
71. A 500-W industrial carbon dioxide cutting laser is capable of cutting 3.0-mm thick steel at a rate of 20 mm/s. Carbon dioxide lasers operate at a wavelength of 10.6 μm. If the spot size of the cutting laser is 0.17 mm in diameter and the intensity is uniform in this spot, what are the magnitudes of the Poynting vector, the rms magnetic field, and the rms electric field for this laser? ••
72. A radio station sends out its signal via a 20-kW emitting antenna. Assuming the signal is emitted equally in all directions and no signal strength is lost, how much power per unit area remains in the signal at a location 10 km from the emitting antenna? ••
73. It takes sunlight about 8 min to travel from the Sun to Earth, where it has an average intensity of 1400 W/m^2. If it takes 44 min for light to travel from the Sun to Jupiter, what is the sunlight intensity there? ••
74. The intensity of sunlight striking Earth's upper atmosphere is approximately 1.35×10^3 W/m^2. (*a*) Calculate E_{max} and B_{max} at this location. (*b*) Given that Earth is 1.5×10^{11} m from the Sun, what is the power output of the Sun? (*c*) If all the sunlight that strikes the upper atmosphere reached Earth's surface, how many watts of power would reach Earth's surface? (Hint: The Earth–Sun distance is so great that you can assume all the sunlight is perpendicular to the cross section of Earth.) ••
75. At a distance of 50 mm (about the distance from your ear to the center of your brain), the 824.6-MHz microwave radiation from a cell phone has an intensity of 35 W/m^2. The maximum permissible radiation leaking from a microwave oven at the same distance is 10 W/m^2. It occurs to you that you could estimate the average magnetic fields generated by the two devices. You also know that people concerned about the possible effects of cell phone radiation on the brain often use a Bluetooth headset, which has an intensity of 0.080 W/m^2 at a distance of 50 mm. But in this case the cell phone itself is probably near your waist—on a belt or in a pocket or purse. There must be an improvement in intensities and field magnitudes at the center of the brain or people would not do this. While contemplating these things, you realize that there may also be a concern that the wavelength of cell phone radiation is comparable to the size of the skull. ••• CR
76. Some people are concerned about "electromagnetic smog" and install meshes on their windows to absorb electromagnetic waves. A particular product advertises "50-dB attenuation in the range from 10 MHz to 3 GHz," with attenuation defined as

$$\text{attenuation (dB)} = 10 \log\left(\frac{\text{input intensity}}{\text{output intensity}}\right).$$

The meshes (sometimes called *Faraday cages* or *Faraday meshes*) generally work if the diameter of the holes is smaller than the wavelength of the radiation. You naturally wonder what this means in terms of reductions in the wave intensity and field strengths. However, it is also curious that the manufacturer advertises a frequency range, so you begin to think about the maximum mesh diameter allowed and whether a minimum is required. ••• CR



31

Electric Circuits

31.1 The basic circuit
31.2 Current and resistance
31.3 Junctions and multiple loops
31.4 Electric fields in conductors

31.5 Resistance and Ohm's law
31.6 Single-loop circuits
31.7 Multiloop circuits
31.8 Power in electric circuits

CONCEPTS

QUANTITATIVE TOOLS

Electric circuits surround us. We light our work and living spaces, start our cars, communicate with each other, and cook our food with electric circuits. Almost everything that consumes energy in our offices and homes is powered by electricity, whether by batteries or by the electric circuitry of the building.

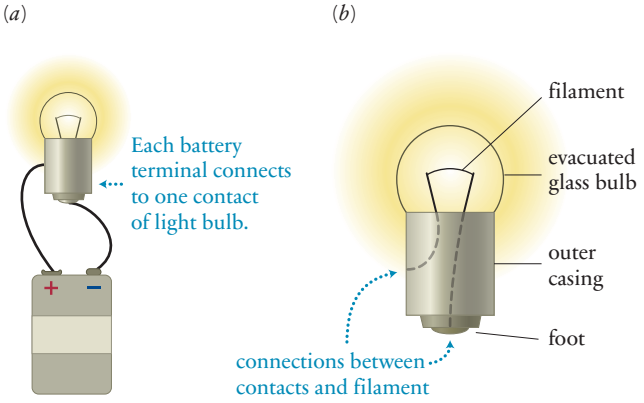
Electrical devices are ubiquitous because electric circuits offer an extremely versatile means of producing and distributing electrical energy for a variety of tasks. Circuits are designed to control the flow of charge carriers—the current—in a device. The current then provides energy to the device, allowing it to perform its function.

In this chapter we explore the basic principles of electric circuits powered by sources of electric potential energy that maintain a constant potential difference, such as batteries. Such circuits are known as *direct-current circuits* or *DC circuits*. In the next chapter, we'll learn about *alternating-current* (*AC*) circuits, which run on time-varying potential differences, such as the electricity delivered to buildings by means of electrical power lines.

31.1 The basic circuit

One familiar way to make electricity do something useful is to connect a battery to a light bulb. To do so, we connect each terminal of the battery to one of the two contacts on the light bulb (**Figure 31.1*a***). On incandescent light bulbs (discussed throughout this chapter), the two contacts are the threaded metal casing and a metallic "foot" that is separated from the casing by an insulator (Figure 31.1*b*). Inside the bulb, each contact is connected by a metal wire to one end of a tungsten *filament*, a very thin wire of tungsten wound in a tight coil. When connected as shown in Figure 31.1, the wires, contacts, and filament form a continuous conducting path from one terminal to the other, and charge carriers flow through the filament. The current in the filament causes the filament to become white hot and glow. (Tungsten is used for light bulb filaments because its very high melting temperature allows it to glow and not melt.)

Figure 31.1 A light bulb connected to a battery.



Experiments with light bulbs show that:

1. A bulb doesn't glow—not even briefly—if only one contact is connected to a battery terminal.
2. In order for a bulb to glow, both contacts must be connected through a continuous conducting path to the terminals of a battery. If the path is broken, the bulb goes out. It doesn't matter where in the path the break occurs.
3. A glowing light bulb generates light and thermal energy.
4. The wires connecting a light bulb to a battery usually do not get hot.
5. A light bulb left connected to a battery over a long time interval eventually goes out because the battery runs down. Once this happens, not enough chemical energy remains in the battery to maintain a large enough potential difference to light the bulb.

In the rest of this chapter we explore the reasons for these observations. The first two observations can be summarized as follows: In order for the light bulb to glow, one of the bulb's contacts must be connected to the positive terminal of the battery and the bulb's other contact must be connected to the negative terminal of the battery. Because the battery maintains a potential difference between its terminals, charge carriers move from one terminal toward the other when a conducting path is provided between the terminals.

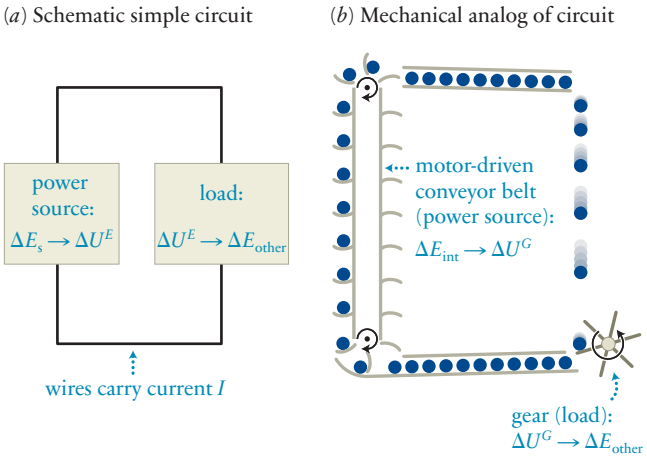
The arrangement shown in Figure 31.1 is an example of an **electric circuit**—an interconnection of electrical components (called *circuit elements*). Any closed conducting path through the circuit is called a **loop.** We shall first study electric circuits that have a single loop.

CONCEPTS

31.1 (*a*) Consider the system comprising the single-loop circuit shown in Figure 31.1*a*, including the light and thermal energy generated by the bulb. (*a*) Is this system closed? (*b*) Is the energy of the system constant? (*c*) Where do the light energy and the thermal energy come from?

Checkpoint 31.1 demonstrates an important feature of circuits: Energy is converted from electric potential energy to some other form. With this in mind, we can draw a general single-loop circuit (**Figure 31.2*a***) consisting of a power source, a load, and wires that connect the load to the source. The **power source** provides electric potential energy to the rest of the circuit, usually by converting some form of energy to electric potential energy. The potential difference across the terminals of the power source drives a current in the circuit. The **load** in an electric circuit is all the circuit elements connected to the power source. In the load, the electric potential energy of the moving charge carriers is converted to other forms of energy, such as thermal or mechanical

Figure 31.2 The energy conversions in a single-loop circuit and in a mechanical analog of the circuit.



energy. The wires connecting the elements in a circuit are considered to be ideal; that is, the wires serve to transport electric potential energy to the load, and we ignore the small amount of energy dissipated in the wires.

Figure 31.2*b* shows a mechanical analog of a single-loop electric circuit. The motor-driven conveyor belt is analogous to the power source in the electric circuit. The conveyor belt lifts the balls, converting source energy to gravitational potential energy. The balls then roll through a horizontal tube, losing only a tiny amount of gravitational potential energy in the process, in the same manner that charge carriers flow with very little loss of energy through wires. Next, the balls drop from the tube onto the gear, which corresponds to the load of an electric circuit. As the balls fall and make the gear turn, their gravitational potential energy is converted to mechanical energy. Finally, the balls roll down a second tube to the bottom of the conveyor belt, completing the cycle.

Notice two things about the mechanical circuit of Figure 31.2*b*. First, because the system is filled with balls, the balls have to move through the circuit one after the other. If the right end of the upper tube were blocked rather than open (the equivalent of the connecting wire in Figure 31.2*a* being disconnected from the load), no more balls could be pushed into it from the conveyor belt. There has to be a closed path through the system in order for the balls to move, just as charge carriers can flow through a circuit only when there is an unbroken conducting path.

The second thing to notice in Figure 31.2*b* is that there is no net transport of balls from the power source to the load. There are as many balls traveling from the power source to the load as the other way around. The balls travel around a closed path through the system, and if we watched one particular ball long enough, we would see it circle repeatedly around the circuit. The balls are simply vehicles transporting energy through the system.

We can therefore propose an operational description of a single-loop DC circuit in terms of energy conversion:

> **In a single-loop DC circuit, electric potential energy acquired by the carriers in the power source is converted to another form of energy in the load.**

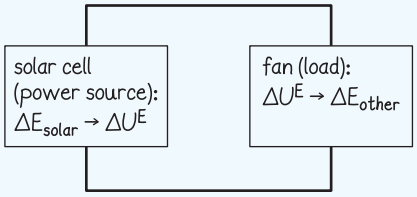
The power source in such an electric circuit doesn't have to be a battery—it can be anything that produces a constant potential difference, thereby driving a current around the circuit, such as a solar cell or an electric generator. Likewise, the load doesn't have to be a light bulb. A toaster acting as a circuit load converts electric potential energy to thermal energy, a loudspeaker acting as a load converts electric potential energy to mechanical energy in sound waves, and a motor load converts electric potential energy to mechanical energy.

EXERCISE 31.1 Solar fan

A solar cell, which converts solar energy (a form of source energy, see Section 7.4) to electric potential energy, is connected to a small fan. Represent this circuit by a diagram analogous to Figure 31.2*a* and describe the energy conversions taking place in this circuit.

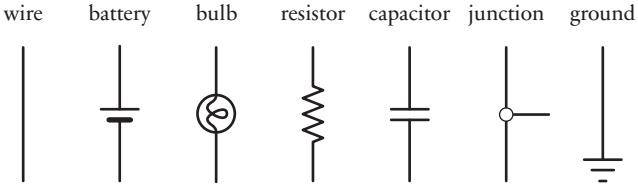
SOLUTION The solar cell is the power source, and the fan is the load. **Figure 31.3** shows my diagram for this circuit. The solar cell converts solar energy to electric potential energy, and the fan converts electric potential energy to another form of energy by setting the air in motion. I add these conversions to my diagram. ✔

Figure 31.3



The various parts of electric circuits are commonly represented by graphical symbols rather than by such words as *power source* and *load*. **Figure 31.4** shows the symbols for some common circuit elements. In

Figure 31.4 Standard representations of common elements of electric circuits.



CONCEPTS

Figure 31.5 Circuit diagram of bulb-and-battery circuit of Figure 31.1.

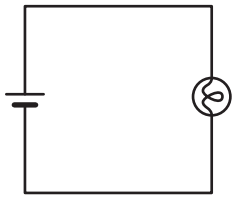


Figure 31.5 these symbols are used to represent the bulb-and-battery circuit of Figure 31.1. Note that in the battery symbol the short, heavy line represents the negative terminal and the longer, thin line represents the positive terminal. A schematic representation of a circuit using the standard symbols shown in Figure 31.4 is called a **circuit diagram.**

31.2 Two wires connect the plates of a charged capacitor to the contacts of a light bulb. (*a*) Does this assembly constitute a circuit? If so, identify the power source and the load. (*b*) Does the bulb glow? (*c*) What energy conversions take place after the bulb and capacitor are connected?

31.2 Current and resistance

Why does the filament of a light bulb get hot and glow, but the wires connecting the bulb to the battery do not? And why does the battery run out of energy and the bulb eventually stop glowing? To answer these questions, we need to look at the motion of charge carriers through an electric circuit.

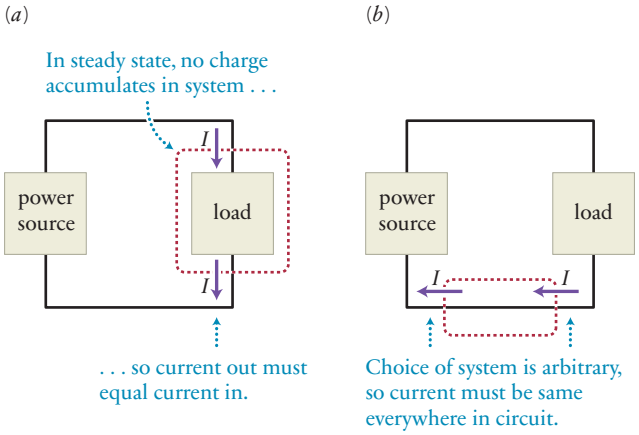
31.3 Suppose you connect a light bulb to a battery. How do you expect the current in the bulb to vary over the course of (*a*) a minute, (*b*) a few days?

CONCEPTS

Experiments show that as long as the power source in a circuit like the one in Figure 31.1 maintains a constant potential difference across its terminals, the current remains constant. Given that a battery connected to a bulb can maintain a constant potential difference for hours, the current in a circuit like the one in Figure 31.1 is constant over that same time interval. This current is established almost instantaneously after the circuit is completed (that is, after all the circuit elements are connected together), and it vanishes almost instantaneously when the circuit is broken. Other than the instants after the circuit is completed or broken, therefore, we have a **steady state** with constant current.

As we found in Section 25.3, positively charged particles tend to move toward regions of lower potential and negatively charged particles tends to move toward regions of higher potential. Indeed, this is why charge carriers flow through a closed circuit. For metal conductors, the mobile charge carriers are electrons; when a metallic conducting path is provided, electrons flow from the negative terminal of the power source to the positive terminal. In materials in which the mobile charge carriers are positively charged, the carriers travel in the opposite direction—from the positive terminal of the power source to the negative terminal. The rate at which the charges on the carriers cross a section in the conductor is the current, and the sign of the carriers and the direction of their motion determine the direction of the current (see Sections 27.3 and 27.5). Remember, however, that for a given potential difference between two points, the current direction is from points of higher potential to points of lower potential regardless of the sign of the mobile charge carriers (see Section 27.3).

Figure 31.6 Because charge is conserved, in steady state it doesn't accumulate in the load or in any other part of the circuit. Hence, in steady state the current into any part of the circuit must be the same as the current out of that part.



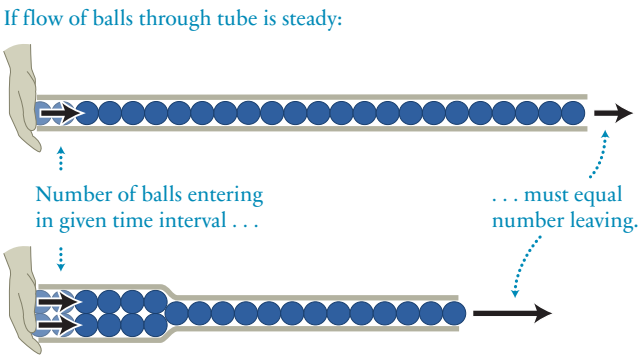
In steady state we can draw an important conclusion regarding currents in circuits. Consider, for example, the system comprising just the load in the circuit shown in **Figure 31.6*a***. Because charge is conserved (see Section 22.3), the amount of charge inside this system can only change due to a flow of charge into or out of the system. In steady state the charge of the system is constant, and so the flow of charge carriers into the system must be equal to the flow of charge carriers out of it. Put differently, the current into the system must equal the current out of the system.

Because we can choose any system we like—as in Figure 31.6*b*, for example—we conclude:

> **In a steady state, the current is the same at all locations along a single-loop electric circuit.**

We shall refer to this requirement as the **current continuity principle.** **Figure 31.7** illustrates the principle for balls flowing through a tube. Because the tube is filled, if the flow of balls through the tube is steady, there is nowhere for the balls to pile up without changing the rate of flow. As a result, when one ball is pushed into the left end of the tube, one ball must come out at the right end. Likewise, if two balls are pushed into the left end at the same instant, two must come out at the right end.

Figure 31.7 The continuity principle.



The current continuity principle tells us something else about the operation of circuits: As electrons flow through a light bulb, they are not accumulating, or "used up," in the bulb. For every electron that goes into one end of the filament, one electron comes out the other end. What then is consumed (used up) by the bulb as it glows? In our discussion of the mechanical circuit of Figure 31.2*b*, we found that the balls moving through the circuit are not consumed anywhere. Instead, they act to transfer energy from the power source to the load. Before we examine the energy of the electrons in the electric circuit of Figure 31.2*a*, answer the next checkpoint.

31.4 Does an electron lose or gain electric potential energy (*a*) while moving inside a battery from the positive terminal to the negative terminal and (*b*) while moving through the rest of the circuit from the negative battery terminal to the positive terminal? (*c*) While flowing through the wire and load portions of the circuit, where do the electrons lose most of their energy?

As we saw in Section 26.4, *inside* a battery, chemical energy is converted to electric potential energy, and electrons at the negative terminal have greater potential energy than those at the positive terminal. In the rest of a circuit, electrons lose that same amount of electric potential energy as they move from the negative terminal through the load to the positive terminal, and that energy is converted to other forms of energy.

For a charged particle moving between two locations through the load, the difference in potential energy from one location to the other is equal to the potential difference between those two locations multiplied by the charge of the particle. Therefore, for each electron that moves through the load of a circuit, an amount of electric potential energy eV_{load} is converted to some other form of energy (light, say, or thermal energy, or mechanical energy), where V_{load} is the magnitude of the potential difference across the load.*

*When the subscript of potential refers to a device or circuit element (say, V_{device}), the symbol is taken to denote the *magnitude* of the potential difference across the device or circuit element.

Thus, changes in the potential difference in an electric circuit show us where energy conversion is taking place. Experiments indicate that V_{load} is essentially equal to V_{batt}, while V_{wire} is negligible. In other words, essentially all of the energy conversion in the circuit takes place in the load and the source, and almost none takes place in the wires.

Because we can ignore the energy dissipation that takes place in the wires:

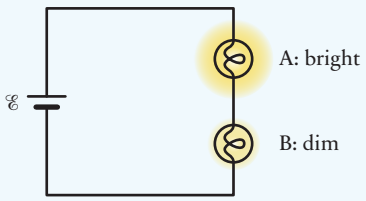
Every point on any given wire is essentially at the same potential.

Therefore we can consider the load in a circuit to be connected directly to the battery if the two are connected to each other through wires.

EXAMPLE 31.2 Current and potential difference

In **Figure 31.8**, two light bulbs are connected to each other by a wire and the combination is connected to a battery. In steady state, bulb A glows brightly and bulb B glows dimly. If the magnitude of the potential difference across the battery terminals is 9 V, what can you say about the magnitude of the potential difference across A?

Figure 31.8 Example 31.2.



1 GETTING STARTED The battery and light bulbs constitute a circuit in steady state. The electrons in this circuit gain electric potential energy inside the battery, and this electric potential energy is converted to light (a form of energy) and thermal energy in the bulbs. I assume that the ratio of light to thermal energy is the same in the two bulbs. Because bulb A glows more brightly than bulb B, bulb A must convert more electric potential energy than bulb B.

2 DEVISE PLAN Because this circuit is in steady state, I know that the current is the same throughout the circuit, and therefore during any given time interval, the same number of electrons pass through both bulbs. For each electron that moves through the load of the circuit, an amount of electric potential energy eV_{load} is converted to light and thermal energy. Therefore the magnitude of the potential difference across each bulb is proportional to the amount of electric potential energy converted to light and thermal energy.

3 EXECUTE PLAN The fact that A glows more brightly than B tells me that A is producing more light and thermal energy. Bulb A must therefore be converting more electric potential energy to these other forms, and the electrons must be losing more energy in A than in B. Because the electrons lose more energy in A than in B, the magnitude of

(Continued)

CONCEPTS

the potential difference across A, V_A, must be greater than the magnitude of the potential difference across B, V_B. Because $V_A > V_B$ and because $V_A + V_B$ must equal V_{batt}, it must be true that V_A is more than half of V_{batt}, which means $9.0\ V > V_A > 4.5\ V$. ✔

4 EVALUATE RESULT It makes sense that the bulb that has the greater potential difference across it glows more brightly.

Example 31.2 illustrates an important point: The magnitudes of the potential difference across various circuit elements in an electric circuit need not be the same, even if the current in them is the same.

The two light bulbs in Figure 31.8 are said to be connected *in series:* There is only a single current path through them, and the charge carriers flow first through one and then through the other. As we have seen in Example 31.2:

The potential difference across circuit elements connected in series is equal to the sum of the individual potential differences across each circuit element.

Experiments show that to obtain a particular current in a circuit element, different elements require widely different potential differences. As noted in passing in Section 29.5, for a given circuit element, the constant of proportionality between the potential difference across the element and the current in it is called the **resistance** of the circuit element:

The resistance of any element in an electric circuit is a measure of the potential difference across that element for a given current in it.

The greater the potential difference required to obtain a certain current in a circuit element, the greater the element's resistance.

The resistance of a particular circuit element depends on the material of which it is made, its dimensions, and its temperature. The resistance of a combination of circuit elements connected in series is equal to the sum of the individual resistances of each element. Therefore, the more light bulbs we connect in series to a battery, the greater the resistance of the combination of light bulbs and the smaller the resulting current in the circuit.

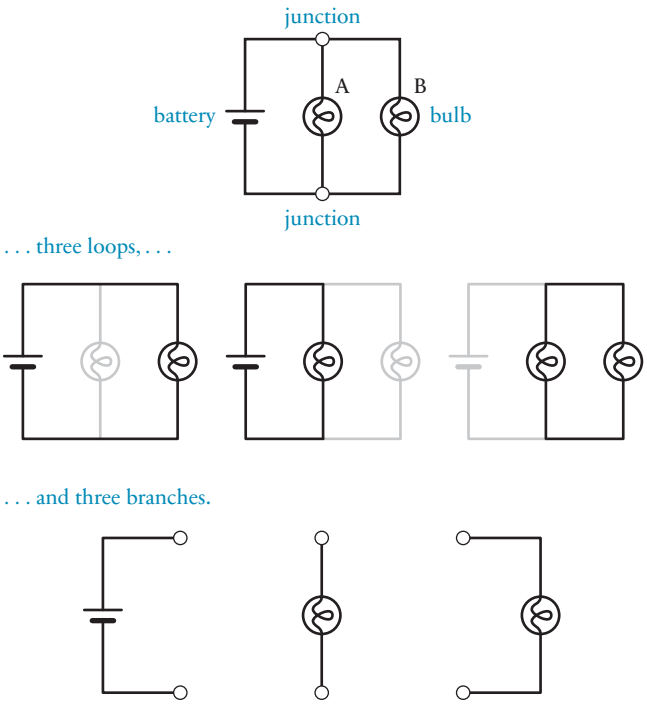
CONCEPTS

31.5 (*a*) Which light bulb in Example 31.2 has the greater resistance? (*b*) Suppose you connect each bulb separately to the battery. Do you expect the current in bulb A to be greater than, equal to, or smaller than that in bulb B?

31.3 Junctions and multiple loops

Circuit elements can also be connected so that more than one conducting path is formed, as illustrated in Figure 31.9. Such circuits contain more than one loop and are called *multiloop circuits*. These circuits contain

Figure 31.9 Circuit diagram of two light bulbs connected to a battery.



junctions—locations where more than two wires are connected together—and **branches**—conducting paths between two junctions that are not intercepted by another junction. The circuit in Figure 31.9, for example, contains two junctions (represented by open circles), three branches, and three loops.

The continuity principle permits us to draw some important conclusions about the currents in multiloop circuits. Let us begin by applying the current continuity principle to the circuit shown in Figure 31.10. Because we can select any system boundary along a branch, the current continuity principle requires the current to be the same throughout that branch. We can apply the same reasoning to the other two branches, so:

The current in each branch of a multiloop circuit is the same throughout that branch.

Figure 31.10 The branch rule: In each branch, the current must be the same throughout that branch.

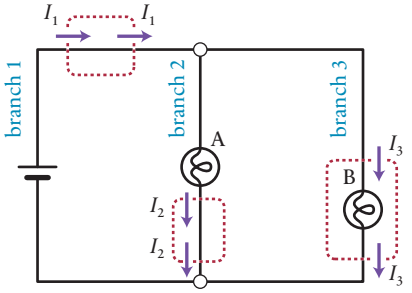
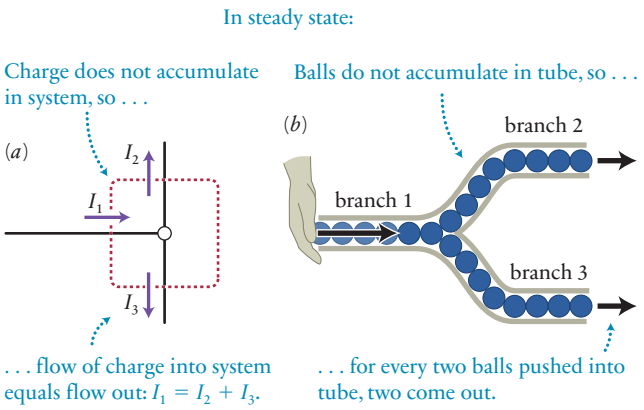


Figure 31.11 The continuity principle at a junction is illustrated by the flow of balls through a branched tube.



This statement is known as the **branch rule.** We shall label the current in each branch according to the branch. In Figure 31.10, for example, the currents in the three branches are I_1, I_2, and I_3.

To examine how charge carriers flow at a junction, we draw a system boundary around a junction, as in **Figure 31.11a**. Because the charge inside the system is not changing and because charge is conserved, it follows that the flow of charge carriers into the system must be equal to the flow of charge carriers out of the system. Specifically, if a current I_1 goes into the junction and the other two wires carry currents I_2 and I_3 out of the junction, then $I_2 + I_3$ must equal I_1. The general statement of this principle is known as the **junction rule:**

> **The sum of the currents directed into a junction equals the sum of the currents directed out of the same junction.**

Figure 31.11*b* illustrates this principle for balls flowing through a junction of tubes. If we push four balls in at one end of branch 1, four balls have to come out from branches 2 and 3 combined. In the junction shown in the figure, branch 2 and branch 3 are equivalent and two balls are likely to come out from each. In general, however, the branches do not need to be equivalent, and in that case the number of balls going into branch 1 must be equal to the sum of the balls coming out of branches 2 and 3. Because the pushing in and the coming out occur during the same time interval, the "current" of balls through branch 1 equals the sum of the currents in branches 2 and 3, which is what we concluded for the currents in Figure 31.11*a*.

Let us return now to the circuit in Figure 31.9. Because the two bulbs are each connected to the same two junctions, they are said to be connected *in parallel* across the battery. Because each junction is at a given potential, we can conclude that for parallel circuit elements:

> **The potential differences across circuit elements connected in parallel are always equal.**

More specifically, for the circuit shown in Figure 31.9, the potential difference across the two light bulbs is equal to the emf of the battery.

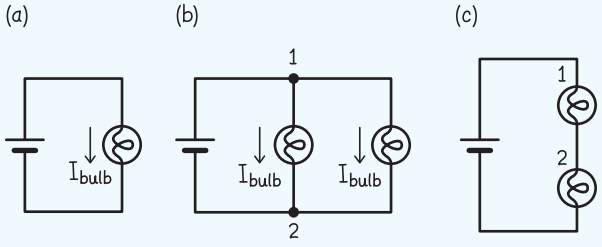
In the next example we study the resistance of parallel combinations of circuit elements.

EXAMPLE 31.3 Series versus parallel

Two identical light bulbs can be connected in parallel or in series to a battery to form a closed circuit. How do the magnitudes of the potential differences across each bulb and the current in the battery compare with those in a single-bulb circuit when the bulbs in the two-bulb circuit are connected in parallel? When they are connected in series?

❶ **GETTING STARTED** I begin by drawing circuit diagrams for the single-bulb circuit and for the parallel and series two-bulb circuits (**Figure 31.12**). To connect the bulbs in parallel to the battery, one contact of each bulb is connected directly to the positive terminal of the battery, and the other contact of each bulb is connected directly to the negative terminal of the battery. For the series circuit, one contact of bulb 1 is connected directly to the positive terminal of the battery, the other contact of bulb 1 is connected to one contact of bulb 2, and the other contact of bulb 2 is connected directly to the negative terminal of the battery.

Figure 31.12



❷ **DEVISE PLAN** To determine the magnitude of the potential difference across each bulb, I need to analyze how the bulbs are connected to the battery. To determine the current in the battery, I must first determine the current through each bulb. To determine the current in each bulb, I use the fact that the resistance and potential difference determine the amount of current in each bulb. I can use the junction rule to determine the currents in the parallel circuit, and so I label the junctions 1 and 2 in Figure 31.12*b*. Because there is only one branch in the series circuit, I know that the current in the battery is the same as the current in the bulbs. Because the bulbs are identical, their resistances are identical too. I assume the wires have no resistance.

❸ **EXECUTE PLAN** Because in the parallel circuit one contact of each bulb is connected to the battery's positive terminal and the other contact of each bulb is connected to the negative terminal, the potential difference across each bulb is equal to the potential difference across the battery. By the same argument, the potential difference across the bulb in the single-bulb circuit is also equal to the potential difference across the battery. Therefore the potential

(Continued)

CONCEPTS

difference across each bulb in the parallel circuit is the same as that across the bulb in the single-bulb circuit. ✔

Because the potential difference is the same across these three bulbs, the current must also be the same in all three bulbs. I'll denote this current by I_{bulb}. In the parallel circuit, the fact that the current in each bulb is I_{bulb} means that the current in the battery must be $2I_{bulb}$. (The current pathway at junction 1 of Figure 31.12*b* is just like the pathway shown in Figure 31.11, with $I_2 = I_3 = I_{bulb}$.) In the single-bulb circuit, the current must be the same at all locations in the circuit, so the current in the battery must be I_{bulb}. ✔

In the series circuit, the potential difference across the two-bulb combination is equal to the potential difference across the battery, which means the magnitude of the potential difference across each bulb must be half the magnitude of the potential difference across the battery. The potential difference across each bulb is thus equal to half the potential difference across the bulb in the single-bulb circuit. ✔

Because the potential difference across each bulb in series is half the potential difference across the battery, the current in each bulb must be half the current in the single-bulb circuit, $I_{bulb}/2$. Because the series circuit in Figure 31.12*c* is a single-loop circuit, the current is the same at all locations. Therefore the current in the battery must also be $I_{bulb}/2$. The current in the battery in the single-bulb circuit is therefore twice that in the series two-bulb circuit. ✔

In tabular form my results are

	Parallel	Series	Single
V_{bulb}	V_{batt}	$V_{batt}/2$	V_{batt}
I_{batt}	$2I_{bulb}$	$I_{bulb}/2$	I_{bulb}

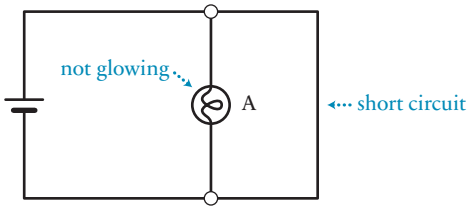
4 EVALUATE RESULT The current in the battery in the parallel circuit is therefore four times that in the series circuit. That makes sense because in the parallel circuit each bulb glows identically to the bulb in the single-bulb circuit, while in the series circuit, the battery has to "push" the charge carriers through twice as much resistance. Therefore it makes sense that in the series circuit, both the potential difference across each bulb and the current in the battery are much smaller than they are in the parallel circuit.

CONCEPTS

31.6 In Figure 31.9, treat the parallel combination of two light bulbs as a single circuit element. Is the resistance of this element greater than, equal to, or smaller than the resistance of either bulb?

Checkpoint 31.6 highlights an important point about electric circuits: Adding circuit elements in parallel *lowers* the combined resistance and *increases* the current. How can adding elements with a certain resistance lower the combined resistance? The resolution to this apparent contradiction is that adding elements in parallel really amounts to adding paths through which charge carriers can flow, rather than adding resistance to existing paths in the circuit. If you are emptying the water out of a swimming pool, it empties faster if you have multiple hoses draining the water than if you drain the entire pool through a single hose.

Figure 31.13 A wire that is connected to a battery *in parallel* to a light bulb constitutes a short circuit.



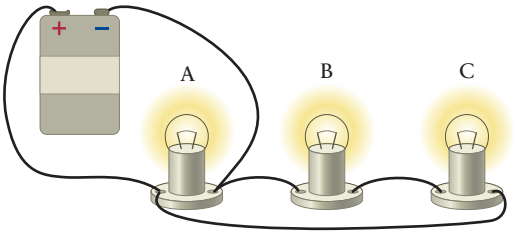
Instead of connecting two light bulbs in parallel as in Figure 31.9, what if we connected a wire in parallel with a bulb as in **Figure 31.13**? Experimentally, we find that replacing bulb B of Figure 31.9 with a wire causes bulb A to stop glowing. Why does this happen? The potential difference across a branch is determined by the potential difference between the two junctions on either end of the branch, and therefore the potential difference across all branches between two junctions must be the same. Because the wire is made from a conducting material, the potential difference between its two ends is zero, and therefore the two junctions in Figure 31.13 are at the same potential. Consequently there is no current in the light bulb.

That doesn't mean there is no current in the circuit. As we'll discuss in more detail in Section 31.4, because of the wire's very small resistance, there is a very large current in the wire. Therefore, if the wire is left connected to the battery as in Figure 31.13, the battery quickly discharges through the wire and the wire heats up. A circuit branch with negligible resistance in parallel with an element is commonly called either a **short** or a **short circuit.**

The circuits in Figures 31.8, 31.10, and 31.12 are all drawn neatly with the circuit elements on a rectangular grid. Real circuits are rarely so neatly laid out, of course, and typically look more like the one shown in **Figure 31.14**. It takes time and concentration to look at the tangle of wires in Figure 31.14 and figure out whether or not the bulbs light up (they do) and identify the path taken by the charge carriers through the circuit.

To analyze such a circuit, therefore, it is helpful to draw a circuit diagram, with elements and wires arranged

Figure 31.14 Real circuits don't always look like circuit diagrams.

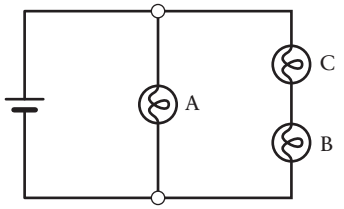


horizontally and vertically. The circuit diagram must accurately show the connections between elements that are present in the actual circuit, but wires that are not connected to each other should, as far as possible, be drawn so that they do not cross.

To draw a circuit diagram for the circuit in Figure 31.14, we begin by identifying the junctions and the branches connecting the junctions. Note that the two terminals of bulb A are connected to two wires each. Because the terminals of the bulb are also connected to each other via the filament inside the bulb, each of these terminals is a junction. There are three branches. One branch consists of the battery and the wires that connect the left and right contacts of light bulb A. The second branch connects the left contact of light bulb A through the filament to the right contact of light bulb A. The third branch connects the right contact of light bulb A through light bulbs B and C to the left contact of light bulb A. **Figure 31.15** shows a circuit diagram for the circuit in Figure 31.14. Junctions are marked by open circles. The battery and light bulbs have been replaced by the symbols introduced in Figure 31.4. The connecting wires between the circuit elements have been straightened out and replaced by lines.

It is not important that the branch of the circuit containing bulb A appears to the left of the branch containing bulbs B and C. The two-bulb branch could be shown to the left of the branch containing bulb A, and the diagram would still be correct. It is also not important that the branch that contains the battery is shown at the left. There is always more than one way to draw a circuit diagram for any given circuit; any diagram that correctly represents the connections between elements in the circuit is valid.

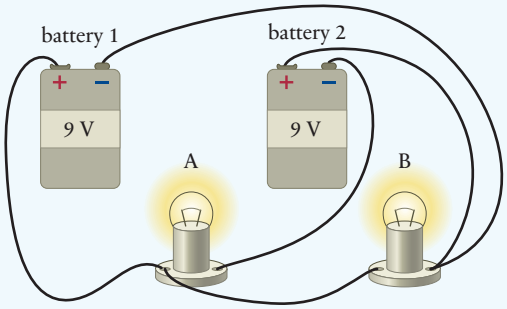
Figure 31.15 Circuit diagram for the circuit of Figure 31.14.



EXERCISE 31.4 Bulbs and batteries

Draw a circuit diagram for the arrangement shown in **Figure 31.16**.

Figure 31.16 Exercise 31.4.

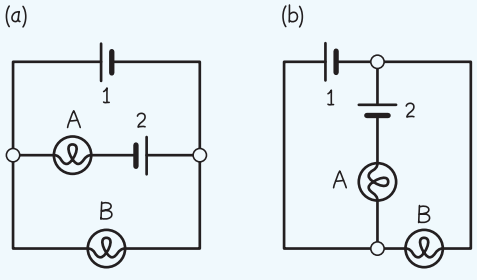


SOLUTION I begin by identifying the junctions and branches. I identify two junctions: one is the left terminal of bulb A and the other is the right terminal of bulb B. I can trace out three branches going from the left terminal of bulb A to the right terminal of bulb B. One branch includes battery 1. The second branch includes bulb A and battery 2, and the third branch includes just bulb B.

I begin by drawing the two junctions. Then I connect these two junctions by each of the three branches, making sure to get the directions of the batteries correct—positive terminal of battery 1 connected to the left terminals of bulbs A and B, and negative terminal of battery 1 connected to the positive terminal of battery 2 and bulb B. This yields the diagram shown in **Figure 31.17*a***.

Note that the diagram I drew is not unique. Had I chosen to draw one junction above the other, I might have obtained the diagram shown in Figure 31.17*b*. By comparing the two diagrams, I see that by sliding battery 1 in Figure 31.17*b* to the leftmost vertical branch, sliding bulb B to the rightmost branch, and then rotating the diagram 90° clockwise, I obtain the diagram shown in Figure 31.17*a*.

Figure 31.17



31.7 If each battery in Figure 31.16 has an emf of 9 V, what is the magnitude of the potential difference across (*a*) bulb A and (*b*) bulb B? (*c*) If the two bulbs are identical, which one glows more brightly?

CONCEPTS

31.4 Electric fields in conductors

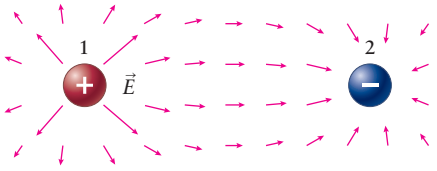
In Section 24.5 we found that the electric field is zero inside a conducting object in electrostatic equilibrium. A conductor through which charge carriers flow is *not* in electrostatic equilibrium, however. (Remember, *electrostatic* means that the arrangement of charged particles is fixed.) To keep charge carriers flowing through a conductor, we need an electric field inside it.

Let us examine more closely the electric field in current-carrying conductors. Consider connecting the two charged spheres of **Figure 31.18*a*** with a metal rod that is much thinner than the radius of the spheres. Charge carriers can now flow from one sphere to the other along this rod (Figure 31.18*b*). For simplicity, we assume that a power source (not shown in the figure) keeps the charge on each sphere and the potential difference between the two constant.

Initially, the electric field along the axis of the rod points from the positive sphere to the negative sphere. However, the electric field is stronger near the spheres, at A and C in Figure 31.18*b*, than in the middle, at B. Consequently, the charge carriers at A and C get pushed along horizontally more strongly than in the middle, causing positive charge carriers to accumulate between A and B and negative carriers to accumulate between B and C. This accumulation changes the electric field in the rod. For example, as charge carriers accumulate between A and B, the flow of carriers at A is reduced and that at B is enhanced. The carriers stop accumulating when the electric field is the same throughout the rod (Figure 31.18*c*). Thus, any unevenness in the electric field causes charge to accumulate, and the accumulation, in turn, affects the electric field. This feedback continues until the electric field is uniform and points along the conductor everywhere. In practice, this evening out takes place in an extremely short time interval (10^{-9} s).

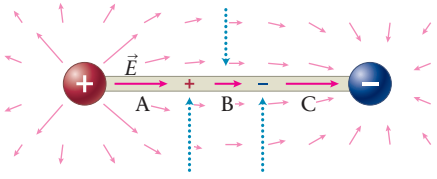
Figure 31.18 The source of the electric field in a current-carrying conductor. (For simplicity, we assume that the charge on each sphere is kept constant by an unseen power source.)

(*a*) Electric field around a pair of charged spheres



(*b*) Just after spheres are joined by conducting rod

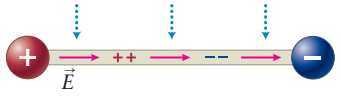
Electric field in rod is smallest in middle (at B).



Unequal electric field in rod causes charge to accumulate.

(*c*) Steady state

Accumulation continues until field is equal throughout rod.

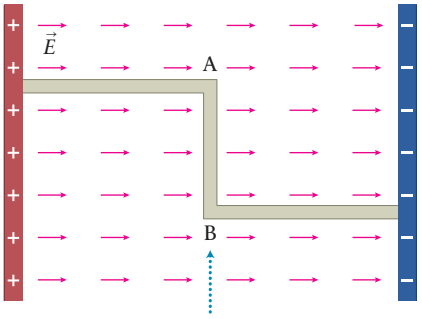


31.8 Suppose the distance between the spheres in Figure 31.18 is ℓ and the potential difference between them is V_{12}. What is the magnitude of the electric field inside the connecting rod?

What is the electric field in a bent conductor, as opposed to the straight conductor of Figure 31.18? Consider the conductor connecting the charged plates in **Figure 31.19*a***. When the plates are first connected to each other via a long thin conductor, the only electric field present is that of the plates, and this field pushes electrons everywhere to the left. Because of this, electrons pile up on the left surface of vertical segment AB of the conductors, leaving behind positive ions on the right side (Figure 31.19*b*). The surface of the conductor

Figure 31.19 Electric field in a bent conductor that connects two charged plates.

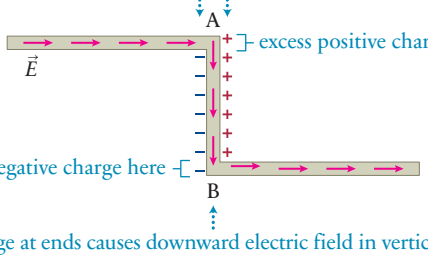
(*a*) Just after charged plates are joined by bent conductor



Electric field of capacitor pushes electrons in vertical segment to left.

(*b*) Steady state

Electric field causes charge to accumulate on surfaces of vertical segment.



Excess charge at ends causes downward electric field in vertical segment.

CONCEPTS

accumulates charge in this manner until the horizontal component of the electric field is zero within segment AB. The electric field within AB then no longer pushes electrons to the left surface.

While charged particles accumulate on the two sides of segment AB, the corners at A and B also acquire a charge (positive at A and negative at B, Figure 31.19*b*). The accumulated charge reduces the electric field in the horizontal portions of the conductor and establishes a downward-pointing electric field in segment AB. Charge accumulates in this manner until the electric field due to the charged plates and the accumulated charge no longer pushes charged particles to the corners, but instead guides them along the conductor.

In general, when the ends of a wire are held at a fixed potential difference, charge accumulates on the surface of the conductor. The accumulation stops when the electric field due to the combined effects of the surface charge distribution and the applied potential difference has the same magnitude everywhere inside the conductor and points along a path through the conductor that is everywhere parallel to the sides of the conductor. Once this electric field is established, it no longer pushes charge carriers to the surface of the conductor but instead drives a steady flow of charge carriers through the conductor:

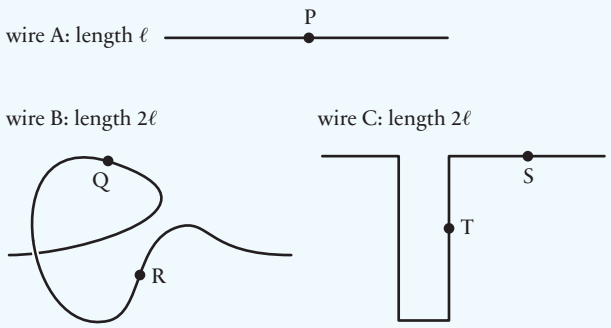
> **In a conductor of uniform cross section carrying a steady current, the electric field has the same magnitude everywhere in the conductor and is parallel to the walls of the conductor.**

Keep in mind that because potential differences across conductors are generally small, the electric fields inside conductors are also generally small, even in the presence of substantial currents.

EXAMPLE 31.5 Bending fields

Consider the three pieces of wire in **Figure 31.20**. All three are made of the same material and have identical circular cross sections. Conductor A has length ℓ; conductors B and C have length 2ℓ. With the wires kept in the configurations shown, a positively charged conducting plate is connected to the left end of each wire and a negatively charged conducting plate is connected to the right end of each wire.

Figure 31.20 Example 31.5.



Rank the magnitudes of the steady-state electric field at P, Q, R, S, and T.

❶ **GETTING STARTED** Once a steady flow of charge carriers has been established in each wire, the electric field in each wire has the same magnitude everywhere in that wire and is always parallel to the walls of the wire.

❷ **DEVISE PLAN** Because the electric field is uniform along each wire, I can use the result I obtained in Example 25.5*a* (Eq. 1): The magnitude of a uniform electric field between two points is equal to the magnitude of the potential difference between those points divided by the distance between those points.

❸ **EXECUTE PLAN** In each case the magnitude of the potential difference between the ends of the conductor is equal to the potential difference between the plates, which I'll denote by V_{plates}. The electric field magnitude in A is then V_{plates}/ℓ, and that in B and C is $V_{\text{plates}}/2\ell$. Therefore, if I use E_P for the electric field magnitude at P, I can say $E_P > E_Q = E_R = E_S = E_T$. ✔

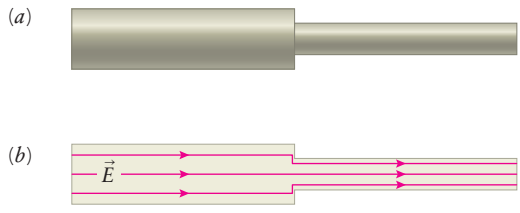
❹ **EVALUATE RESULT** My result is not surprising because I can think of a wire as a load with very small resistance. A wire of length 2ℓ is equivalent to two wires of length ℓ connected in series. Based on the result I obtained in Example 31.3, I expect the current in the wire of length ℓ (analogous to the single-bulb case of Example 31.3) to be twice the current in the wire of length 2ℓ (analogous to the two-bulb series case). To obtain this greater current, I need a greater electric field, as my result shows.

31.9 Sketch the electric field lines inside the conductors in Figures 31.18 and 31.19.

As Example 31.5 shows, the current in wire A is twice the current in wires B and C, which are twice as long. To get the same current in wires B and C as in A, the magnitude of the potential difference across wires B and C must be doubled. This means that the resistance of wires B and C is twice the resistance of wire A. Therefore, the resistance of a conductor is proportional to the length of the conductor.

So far, the only conductors we have considered have all had a uniform cross section. Suppose now that we have a conductor that has a circular cross section that decreases in radius by a factor of 2 at its midpoint (**Figure 31.21a**). How do the current and the electric

Figure 31.21 (*a*) Conductor of circular cross section that decreases in radius by a factor of 2 at the midpoint. (*b*) Electric field lines in the conductor when it is connected between two charged objects (not shown).



CONCEPTS

field in this conductor compare in the wide and the narrow parts?

If we think of this conductor as two circuit elements (the wide part and the narrow part) connected in series, we see that the current must be the same through the entire conductor. At steady state, electric field lines inside a conductor in a circuit must satisfy the continuity principle, just as the current must. Therefore the electric field lines in the wide part of our conductor must all continue into the narrow part (Figure 31.21*b*). The density of the electric field lines, and consequently the magnitude of the electric field, must therefore increase by the same factor by which the area decreases. Because the cross-sectional area decreases by a factor of 4, the magnitude of the electric field must increase by a factor of 4.

Now consider the magnitude of the potential difference across each part of the conductor. In each part, the magnitude of the electric field is equal to the magnitude of the potential difference across the part divided by the length of the part (see Example 31.5). Because the electric field is four times greater in the narrow part and because the two parts have the same length, the potential difference across the narrow part must also be four times greater than the potential difference across the wide part.

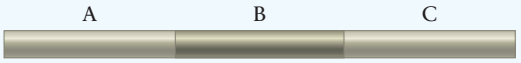
To get the same current in both parts, four times as much potential difference must be applied to the narrow part! This means that the resistance of the narrow part is four times greater than the resistance of the wide part, even though both are made of the same material. Therefore, the resistance of a conductor not only depends on the material from which it is made and is proportional to the length of the conductor, but also is inversely proportional to the cross-sectional area of the conductor.

CONCEPTS

EXAMPLE 31.6 Electric fields in conductors

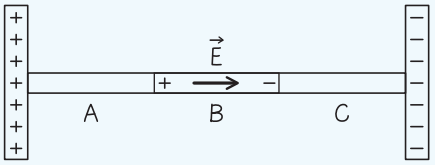
Figure 31.22 shows a rod made of three pieces of conducting material connected end to end. The pieces are of equal size but are made of different materials; pieces A and C have negligible resistance, but piece B has significant resistance. Consider connecting two oppositely charged plates to the ends of this rod. After the plates have been connected, how does the electric field in B compare with the electric field that existed between the plates before they were connected to the rod? Assume the charge on the plates is maintained.

Figure 31.22 Example 31.6.



❶ **GETTING STARTED** I begin by making a sketch of the rod connected to the two charged plates (**Figure 31.23**).

Figure 31.23



❷ **DEVISE PLAN** As I noted in Example 31.5 in using my result from Example 25.5*a*, the magnitude of the electric field in B is equal to the magnitude of the potential difference across B divided by the length of B. Likewise, the electric field magnitude between the plates is equal to the magnitude of the potential difference between the plates divided by the distance between them. To compare these two field magnitudes, I need to compare the magnitude of the potential difference across B with that across the plates.

❸ **EXECUTE PLAN** Because the resistance of A and C is negligible, the potential difference across A and C is negligibly small; the potential difference across B is therefore essentially the same as that between the plates. Because B is one-third as long as the distance between the plates, the electric field magnitude in B is roughly three times the electric field magnitude that existed between the plates before connection. ✔

❹ **EVALUATE RESULT** Because A and C have negligible resistance, I expect the charge to distribute itself along the rod as shown in Figure 31.23, with positive charge residing at the end of B nearer the positively charged plate and negative charge residing at the other end. This distribution of charge reduces the electric field in A and C and increases the electric field in B, as I concluded.

The rod in Figure 31.22 is a common circuit element called a **resistor,** a piece of conducting material that has non-negligible resistance (piece B) attached at both ends to pieces of wire called electrical *leads* (represented in Figure 31.22 by pieces A and C).

31.10 If a potential difference of 9 V is applied across the conductor in Figure 31.21, what is the magnitude of the potential difference (*a*) across the wide part and (*b*) across the narrow part?

Self-quiz

1. Why are none of the bulbs in **Figure 31.24** lit?

Figure 31.24

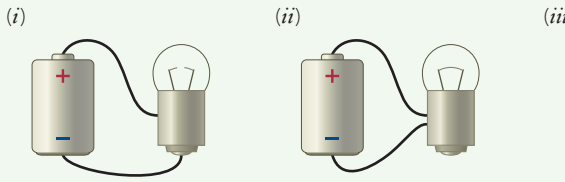
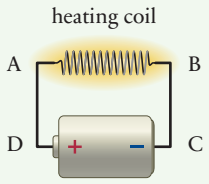


Figure 31.25



2. In **Figure 31.25**, identify the energy conversions that occur between points A and B, B and C, C and D, and D and A.
3. In **Figure 31.26**, bulb B is brighter than bulb C, which in turn is brighter than bulb A. Rank, largest first, (*a*) the magnitudes of the potential differences across the bulbs, (*b*) the currents in them, and (*c*) their resistances.

Figure 31.26

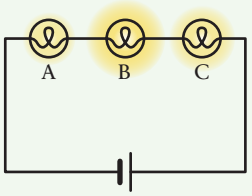


Figure 31.27

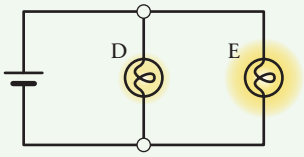
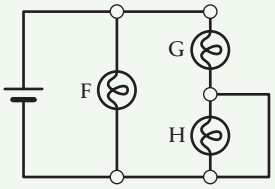
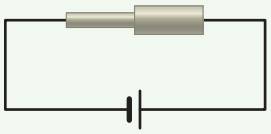


Figure 31.28



4. In **Figure 31.27**, bulb E is brighter than bulb D. Which bulb has (*a*) the greater potential difference across it, (*b*) more current in it, and (*c*) more resistance?
5. In **Figure 31.28**, which of the three bulbs F, G, and H light up?
6. A thick resistor and a thin resistor of the same length and material are connected in series, as shown in **Figure 31.29**. Which resistor has (*a*) the greater potential difference across it and (*b*) the greater resistance?

Figure 31.29

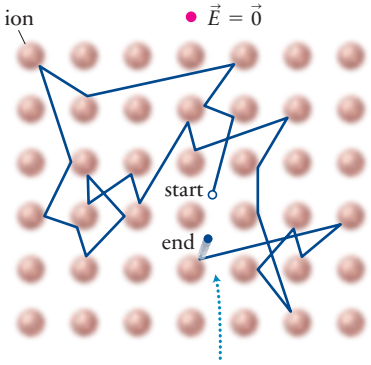


Answers

1. (*i*) Filament broken. (*ii*) Battery connected to only one contact on bulb. (*iii*) Bulb connected to only one battery terminal.
2. AB: electric potential energy converted to thermal energy; CD: chemical energy converted to electric potential energy; BC and DA: negligible conversion of energy.
3. (*a*) $V_B > V_C > V_A$ because the brightest bulb converts the most energy. (*b*) $I_B = I_C = I_A$ because the bulbs are in series and the current is the same throughout the circuit. (*c*) $R_B > R_C > R_A$ because, for a given current in a circuit element, R is proportional to V.
4. (*a*) $V_D = V_E$ because the bulbs are in parallel. (*b*) $I_E > I_D$ because E is brighter. (*c*) $R_D > R_E$ because less current passes through D.
5. F and G light up, but H does not because the wire in parallel with H carries all of the current through the right branch. (The wire "shorts out" H.)
6. (*a*) $V_{thin} > V_{thick}$ because the density of electric field lines through the thin resistor has to be greater, and if the lengths of the resistors are the same and E is greater, then V is greater. (*b*) $R_{thin} > R_{thick}$ because resistance is inversely proportional to cross-sectional area.

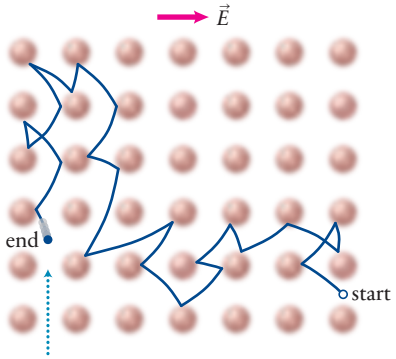
Figure 31.30 The effect of an applied electric field on the motion of a free electron through a lattice of ions.

(*a*) Motion in absence of an electric field



Electron's displacement is very small over long time interval.

(*b*) Motion with applied electric field



Electron undergoes displacement in direction opposite to electric field.

31.5 Resistance and Ohm's law

Recall from Section 22.3 that a metal consists of a lattice of positively charged ions through which electrons can move relatively freely. Although the lattice positions are fixed (meaning there is no motion of ions through the metal), the ions vibrate around their lattice positions, as if they were connected to those positions by springs. (The amplitude of the vibrations depends on temperature.) The ions consist of atomic nuclei surrounded by most of their electrons; the outermost one or more electrons of each atom are free to move through the entire lattice. Because of their thermal energy, these free electrons move through the lattice at very high speeds (10^5 m/s at room temperature, or about 0.1% of the speed of light), and they move in straight lines without any change in energy or momentum except in the instants when they collide with the ions.

In the absence of an electric field and over a time interval long enough for many collisions to take place, the displacement of an electron is very small in spite of its high speed (**Figure 31.30***a*). This is true because each collision changes the direction of the electron's motion, randomizing the direction of the electron's velocity. Consequently, the average velocity of all the electrons is zero. However, when an electric field is applied, as in Figure 31.30*b*, the electric field causes the electrons to accelerate in the direction opposite to the electric field.

A quantitative description of the motion of charge carriers in a conductor was given by P. K. Drude in 1900, shortly after the discovery of the electron. His model (called the *Drude model*) applies remarkably well to metals. Let's describe the motion of the electrons in this model. In the presence of a uniform electric field $\vec{E}$, in the absence of nonelectric interactions, the electrons are subject to a force $-e\vec{E}$ and therefore have an acceleration $\vec{a} = -e\vec{E}/m_e$. For any one electron moving along a straight path between two consecutive collisions, the electron's final velocity $\vec{v}_f$ just before the second collision is

$$\vec{v}_f = \vec{v}_i + \vec{a}\Delta t = \vec{v}_i - \frac{e\vec{E}}{m_e}\Delta t, \tag{31.1}$$

where $\vec{v}_i$ is the electron's initial velocity on that path (its velocity just after the first collision), Δt is the time interval the electron spends on that path (in other words, the time interval between collisions), e is the elementary charge (Eq. 22.3), $\vec{E}$ is the applied electric field, and m_e is the mass of the electron. The magnitude of $\vec{v}_i$ is roughly 10^5 m/s, as noted above, and the direction is determined by the first of the two collisions. Because of the high electron speed, the time interval Δt between collisions is extremely short—on the order of 10^{-14} s.

To calculate the average velocity of all the electrons, we take the average of Eq. 31.1 for all of the electrons:

$$(\vec{v}_f)_{av} = (\vec{v}_i)_{av} - \frac{e\vec{E}}{m_e}(\Delta t)_{av}. \tag{31.2}$$

(I have assumed here that the electric field is either constant over time or takes a time interval much longer than Δt to change significantly.)

Even though the magnitude of $\vec{v}_i$ is quite large, its average value for all electrons is zero because the collisions produce a random distribution of the directions of the initial velocities. The resulting average velocity, called the **drift velocity** $\vec{v}_d$ of the electrons, is thus

$$\vec{v}_d = -\frac{e\vec{E}}{m_e}\tau, \tag{31.3}$$

QUANTITATIVE TOOLS

where $\tau \equiv (\Delta t)_{\text{av}}$ is the average time interval between collisions. (The value of τ depends on the number density, size, and charge of the lattice ions, and on temperature.) The magnitude of the drift velocity is called the *drift speed*. Equation 31.3 shows that the drift velocity of the electrons is in a direction opposite that of the electric field, and the drift speed is proportional to the electric field magnitude.



31.11 (*a*) Does the electric field do work on the electrons of Figure 31.30*b* as they accelerate between collisions? (*b*) On average, does the kinetic energy of the electrons increase as they drift through the lattice? (*c*) What do your answers to parts *a* and *b* imply about the energy in the lattice?

In Chapter 27 we found that the current in a conductor can be expressed in terms of the speed v of the charge carriers (Eq. 27.16). Because the speed of the charge carriers is what we are now calling the drift speed, we can write Eq. 27.16 in the form

$$I = nAqv_{\text{d}}, \tag{31.4}$$

where A is the cross-sectional area of the conductor, q is the charge on each charge carrier in the conductor, and n is the number of carriers per unit volume in the conductor. Because of the current's dependence on cross-sectional area, it is convenient to introduce the current per unit area, called the **current density,** whose magnitude is given by the magnitude of the current per unit area:

$$J \equiv \frac{|I|}{A} = n|q|v_{\text{d}}. \tag{31.5}$$

Because the drift velocity is a vector, the current density is a vector, too:

$$\vec{J} = nq\vec{v}_{\text{d}}. \tag{31.6}$$

The direction of the current density is the same as that of the drift velocity for positive charge carriers and opposite the direction of the drift velocity for negative charge carriers. Therefore the current density is always in the same direction as the current. The SI unit of current density is A/m^2.

Substituting the absolute value of the right side of Eq. 31.3 for v_{d} and e for $|q|$ in Eq. 31.5 yields

$$J = n(e)\left(\frac{eE}{m_{\text{e}}}\tau\right) = \frac{ne^2\tau}{m_{\text{e}}}E. \tag{31.7}$$

Equation 31.7 shows that the current density is proportional to the applied electric field. The proportionality constant σ is called the **conductivity** of the material of which the conductor is made:*

$$\sigma \equiv \frac{J}{E}. \tag{31.8}$$

The SI unit of conductivity is equal to $(\text{A/m}^2)/(\text{V/m}) = \text{A/(V}\cdot\text{m)}$. The conductivity is a measure of a material's ability to conduct a current for a

*Note that conductivity is *not* the same as surface charge density, which is represented by the same symbol.

Table 31.1 Conductivities of various materials at room temperature [A/(V · m)]

Conductors	Silver	6.3×10^7
	Copper	5.9×10^7
	Aluminum	3.6×10^7
	Tungsten	1.8×10^7
	Nichrome	6.7×10^5
	Carbon	7.3×10^4
Semiconductors	Silicon	4×10^{-4}
	Germanium	2
Poor conductors	Seawater	4
Insulators	Pure water	4.0×10^{-6}
	Glass	10^{-12}

given applied electric field. The conductivity is a property of the material and is therefore the same for any piece of that material you might choose.

Table 31.1 gives the conductivities of some common materials at room temperature. Note that the first four materials, all metals, have very similar conductivities. Nichrome is an alloy of nickel and chromium used in heating elements because of its relatively low conductivity. Silicon is a *semiconductor;* its conductivity is intermediate between that of an electrical conductor and that of an electrical insulator. The conductivities of insulators (such as pure water and glass) are many orders of magnitude smaller than those of metals. Seawater has a greater conductivity than pure water because the ions dissolved in seawater serve as charge carriers for a current, just as the ions in the electrolyte of a battery do.

Comparing Eqs. 31.7 and 31.8 gives us an expression for the conductivity of a metal in terms of the average time interval between collisions, the mass and charge of the electrons, and the number density of electrons present in the material:

$$\sigma = \frac{ne^2\tau}{m_e}. \tag{31.9}$$

Because of the temperature dependence of τ, the conductivity σ depends on temperature and, for some materials, on the magnitude of the current in the material, but it is independent of the shape of the piece of material in which the current density is measured.

We obtained Eq. 31.9 for a metal, in which the charge carriers are electrons, but it applies to any system in which the charge carriers in a conducting material move freely between collisions. To generalize Eq. 31.9 for charge carriers other than electrons, simply substitute q^2 for e^2, and m_q for m_e, where q is the charge on the carriers and m_q is the mass of one carrier.

The conductivity describes how large a current density and hence current are created by an external electric field. Such an electric field is produced by applying a potential difference across a material. In Section 31.2, we looked at resistance as a way of relating the applied potential difference V across a circuit element to the resulting current in the element. In general, the resistance of any element is defined to be

$$R \equiv \frac{V}{I}. \tag{31.10}$$

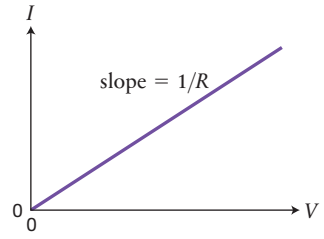
QUANTITATIVE TOOLS

The derived SI unit of resistance is the **ohm** ($1\ \Omega \equiv 1\ \text{V/A}$). Resistance is always positive, so in Eq. 31.10 the direction in which V and I are measured must be such that they both have the same algebraic sign. The resistance of most circuit elements is typically in the range of 10 Ω to 100,000 Ω.

The concept of resistance is useful in describing conductors and other objects that provide a continuous conducting path for charge carriers, such as filaments, bulbs, and resistors. (It is not especially useful for other types of circuit elements, such as capacitors.) For some conducting materials, the resistance R of a piece of this material is fixed at a given temperature. For these materials, Eq. 31.10 indicates that the current in the conductor is proportional to the potential difference across it and inversely proportional to the resistance:

$$I = \frac{V}{R}. \tag{31.11}$$

Figure 31.31 Current versus applied potential difference for an ohmic conductor. The current is proportional to the potential difference.



Such materials are said to be *ohmic*. If we plot the magnitude of the current in a piece of this material as a function of the potential difference across it, the result is a straight line whose slope is equal to the inverse of the resistance R of the piece (**Figure 31.31**).

Equation 31.11 is often referred to as **Ohm's law.** It is important to keep two things in mind about Eq. 31.11. First, many materials and many circuit elements are not ohmic. For such materials and elements, a plot of current as a function of applied potential difference is not a straight line, but has a more complicated shape, indicating that R depends on the potential difference. Second, Eq. 31.11 is really a definition, not a law. It simply amounts to the observation that in certain materials, the current is proportional to the applied potential difference. In this chapter, we concern ourselves only with ohmic materials and circuit elements.

Let's now relate the resistance of a conductor to the conductivity of the material of which it is made. From Checkpoint 31.8 we know that the magnitude of the potential difference across a wire of length ℓ is given by (Eq. 25.25 and Example 25.5):

$$|V| = E\ell. \tag{31.12}$$

Combining Eqs. 31.5 and 31.8, we see that $J \equiv |I|/A = \sigma E$, so $E = |I|/(\sigma A)$. Substituting this expression into Eq. 31.12 and dropping the absolute-value symbol, we obtain a relationship between V and I:

$$V = \frac{I}{\sigma A}\ell = I\frac{\ell}{\sigma A}. \tag{31.13}$$

Substituting this expression into the definition of resistance, Eq. 31.10, we obtain for ohmic conductors,

$$R = \frac{\ell}{\sigma A}. \tag{31.14}$$

Equation 31.14 shows that the resistance of a conductor not only depends on the material from which it is made (through the conductivity σ) but also is proportional to the length ℓ of the conductor and inversely proportional to the cross-sectional area, as we had concluded in Section 31.4.

EXAMPLE 31.7 Drifting electrons

Consider a piece of copper wire that is 10 m long and has a diameter of 1.0 mm. The number density of free electrons in copper is 8.4×10^{28} electrons/m^3. If the wire carries a current of 2.0 A, what are (*a*) the magnitude of the potential difference across the wire, (*b*) the drift speed of the electrons in the wire, and (*c*) the average time interval between collisions for the electrons?

❶ **GETTING STARTED** Ohm's law relates the potential difference across the wire to the current in it and its resistance. Even though the wire is made of copper, it has a finite resistance, which depends on the conductivity of copper and on the length and cross-sectional area of the wire. This resistance is caused by collisions between the electrons and copper ions in the wire, and these collisions limit the drift speed of the electrons through the wire. The drift speed is related to the average time interval between collisions, which is one of the parameters I need to calculate and which appears in the expression for conductivity in the Drude model, Eq. 31.3, $v_d = eE\tau/m_e$.

❷ **DEVISE PLAN** The potential difference across the wire is related to the current by Eq. 31.11, but in order to use this equation, I must determine the wire's resistance, which is given by Eq. 31.14. I can look up the conductivity of copper in Table 31.1 ($\sigma = 5.9 \times 10^7$ A/V·m). To obtain the drift speed of the electrons, I can use Eq. 31.4, which contains the charge on the electron ($q = e = 1.6 \times 10^{-19}$ C), the current (given), the number density of the electrons (given), and the cross-sectional area of the wire, which I can calculate from the given diameter. To obtain the average time interval between collisions, I can use Eq. 31.9, which contains the mass of the electron ($m_e = 9.11 \times 10^{-31}$ kg).

❸ **EXECUTE PLAN** (*a*) I obtain the resistance of the wire from Eq. 31.14 and the cross-sectional area of the wire, $A = \pi r^2 = \pi(5.0 \times 10^{-4}\ \text{m})^2 = 7.9 \times 10^{-7}\ \text{m}^2$:

$$R_{\text{wire}} = \frac{\ell}{\sigma A} = \frac{10\ \text{m}}{(5.9 \times 10^7\ \text{A/V·m})(7.9 \times 10^{-7}\ \text{m}^2)}$$

$$= 0.21\ \Omega.$$

Now that I have R_{wire}, I can obtain V_{wire} from Eq. 31.11:

$$V_{\text{wire}} = IR_{\text{wire}} = (2.0\ \text{A})(0.21\ \Omega) = 0.42\ \text{V}. \checkmark$$

(*b*) I first obtain the drift speed of the electrons from Eq. 31.4:

$$v_d = \frac{I}{neA} = \frac{2.0\ \text{C/s}}{(8.4 \times 10^{28}\ \text{m}^{-3})(1.6 \times 10^{-19}\ \text{C})(7.9 \times 10^{-7}\ \text{m}^2)}$$

$$= 1.9 \times 10^{-4}\ \text{m/s}. \checkmark$$

(*c*) Solving Eq. 31.9 for τ, I get

$$\tau = \frac{m_e\sigma}{ne^2} = \frac{(9.11 \times 10^{-31}\ \text{kg})(5.9 \times 10^7\ \text{A/V·m})}{(8.4 \times 10^{28}\ \text{m}^{-3})(1.6 \times 10^{-19}\ \text{C})^2}$$

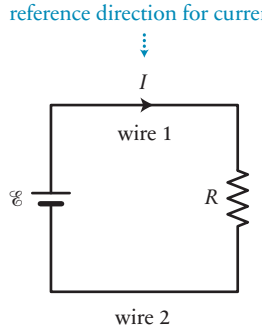
$$= 2.5 \times 10^{-14}\ \text{s}. \checkmark$$

❹ **EVALUATE RESULT** Because the conductivity of copper is high, it makes sense that I obtain a small resistance and consequently a small potential difference across the wire even though it is 10 m long. My answer to part *b* shows that the drift speed of the electrons in the wire is very small. The drift speed indicates that it takes the electrons about 5 s to move just 1 mm, even though I know from experience that when I turn on a light switch, the light turns on instantly even if the bulb is meters away from the switch. The reason the light bulb comes on almost instantaneously, however, is that the current is the same throughout the circuit—all of the electrons throughout the circuit are set in motion almost simultaneously when I flip the switch.

How can my very small calculated value for drift speed be reasonable when the current is a significant 2.0 A = 2.0 C/s at any given location in the wire? The reason is that the number density of electrons in the wire is very high. Although it takes 5 s for an electron to move 1 mm, there are nearly 10^{20} electrons per cubic millimeter of the wire. This means that on the order of 10^{20} electrons, or 10 coulombs of charge, pass a given location in 5 s—hence the current is 2 C/s, or 2 A.

The very high number density of the electrons is also responsible for the extremely short time interval between collisions. If I imagine the 10^{20} electrons per cubic millimeter arranged on a cubic lattice where each cube of the lattice has a side length of 1.0 mm, there are about 10^7 electrons along each side of the cube, and so the average distance between them is about 10^{-10} m. As the electrons move at about 10^5 m/s (see Section 31.5), they cover this average distance in about 10^{-15} s, which is within an order of magnitude of what I obtained.

Figure 31.32 Circuit diagram for a battery connected in series with a single resistor.



31.12 If the temperature of a metal is raised, the amplitude of the vibrations of the metal-lattice ions increases. (*a*) What effect, if any, do you expect these greater vibrations to have on the resistance of a piece of that metal? (*b*) What effect does running a current in a metal have on the temperature of the metal? (*c*) Make a graph of current versus potential difference, taking into account the effect you described in part *b*.

31.6 Single-loop circuits

In this section, we quantitatively analyze single-loop circuits. We begin with the simplest possible circuit: a single battery connected to a single resistor, obeying Ohm's law (Figure 31.32). The emf of the battery is $\mathscr{E}$, and the resistance of the resistor is R.

QUANTITATIVE TOOLS

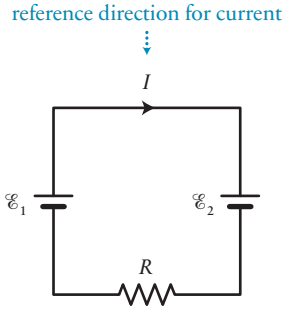
The emf of the battery establishes an electrostatic field that drives a current in the circuit. In steady state, this current is the same at all locations in the circuit, as we discussed in the first part of this chapter. For a single-loop circuit we can express this condition quantitatively as

$$I_{\text{battery}} = I_{\text{wire}} = I_{\text{resistor}} = I. \tag{31.15}$$

In Section 27.3 we called the direction of decreasing potential in a conducting object the *direction of current*, and we indicated this direction by an arrow labeled I next to the object carrying the current (Figure 27.18). When working with circuits, however, we generally won't know the direction of current in advance. To analyze the circuit we therefore need to choose a *reference direction for the current* in each circuit branch. We indicate this direction by an arrowhead on a wire in the circuit (Figure 31.32) and label that arrowhead with a symbol for the current in that branch. In a single-loop circuit, there are no junctions and therefore only one current. As the next checkpoint shows, the reference direction for the current and the direction of current need not be the same. As we shall see later, we obtain a positive value for I when the direction of current and the reference direction for the current are the same. When these two directions are opposite, we find that $I < 0$.

31.13 If $\mathscr{E}_1 < \mathscr{E}_2$ in **Figure 31.33**, is the direction of current the same as the reference direction for the current indicated in the diagram?

Figure 31.33 Checkpoint 31.13.



Let's next consider the energy transformations that occur in the circuit shown in **Figure 31.34**. Nonelectrostatic work done on the negative charge carriers as they travel from point a to point b through the source raises the electric potential energy of the charge carriers. In the load, this electric potential energy is converted to other forms of energy, such as thermal energy, mechanical energy, or radiation energy:

$$\Delta E_{\text{other}} = W_{\text{nonelectrostatic}}(\text{a} \to \text{b}). \tag{31.16}$$

Considering just the load by itself, we also know that the amount of energy converted to other forms of energy must be equal to the work done by the electrostatic field on the charge carriers as they travel from point b to point a through the load:

$$\Delta E_{\text{other}} = W_{\text{electrostatic}}(\text{b} \to \text{a}), \tag{31.17}$$

so

$$W_{\text{nonelectrostatic}}(\text{a} \to \text{b}) = W_{\text{electrostatic}}(\text{b} \to \text{a}). \tag{31.18}$$

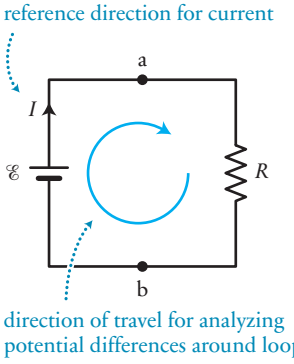
Using Eqs. 26.7 and 25.15 this becomes

$$q\mathscr{E} = -qV_{\text{ba}} \tag{31.19}$$

or

$$\mathscr{E} + V_{\text{ba}} = 0. \tag{31.20}$$

Figure 31.34 Choosing a direction of travel for analyzing a single-loop circuit.



For circuits that contain many elements, Eq. 31.20 can be generalized by replacing each term by a sum. In that case the algebraic sum of the emfs and the potential differences around the loop is zero:

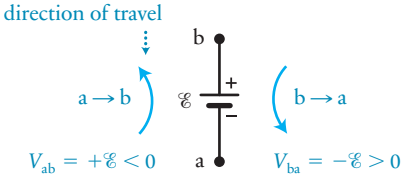
$$\Sigma\mathscr{E} + \Sigma V = 0 \quad (\text{steady state, around loop}). \tag{31.21}$$

Equation 31.21 is called the **loop rule.**

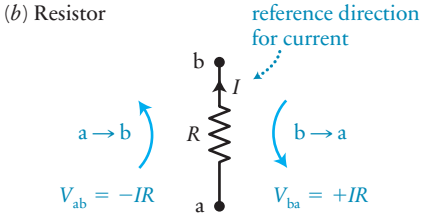
When evaluating the sum on the left in Eq. 31.21, we need to pay close attention to the signs of the emfs and potential differences. We begin by

Figure 31.35 Potential differences across batteries, resistors, and capacitors.

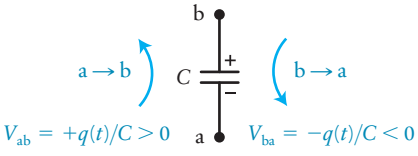
(*a*) Battery



(*b*) Resistor



(*c*) Capacitor



choosing a *direction of travel* around the loop. We denote this direction of travel by a curved arrow at the center of the loop (Figure 31.34). As we shall see shortly, this choice does not affect the end result—we must just be sure to use the same direction for all elements in the loop. In single-loop circuits it makes sense to let the travel direction be the same as the reference direction for the current. In multiloop circuits, however, it is generally not possible to let both directions coincide.

Because the potential is the same everywhere along any of the wires, we only need to consider the potential differences across circuit elements: Each circuit element contributes one term to the sum. For a battery, the potential difference is positive when traveling from the negative terminal to the positive terminal because the potential of the positive terminal is higher than that of the negative terminal (**Figure 31.35*a***). Conversely, the potential difference is negative when traveling in the opposite direction. If the battery is ideal, the magnitude of the potential difference is equal to the emf $\mathscr{E}$ of the battery (Eq. 26.8).

The sign of the potential difference across a resistor depends on both the choice of current reference direction and the choice of travel direction (Figure 31.35*b*). When traveling in the same direction as the reference direction for the current, the potential difference is $-IR$. To see why this is so, let us assume that in Figure 31.35*b* the current is positive (that is, the current is in the reference direction for the current). Because the direction of current is from higher potential to lower potential, we see that the potential at b must be lower than at a, $V_b < V_a$, and therefore $V_{ab} = V_b - V_a < 0$. Indeed when $I > 0$ the potential difference $V_{ab} = -IR$ is negative. When traveling in the opposite direction, the potential difference is $+IR$.

Next we examine the potential difference across a capacitor (Figure 31.35*c*). The potential difference is positive when traveling from the negatively charged plate to the positively charged one because the potential of the positive terminal is higher than that of the negative terminal; the potential difference is negative when traveling in the opposite direction. The magnitude of the potential difference can be obtained from Eq. 26.1: $C \equiv q/V_{cap}$, where C is the capacitance and q is the magnitude of the charge on the plates. It is important to keep in mind that the charge on the plate and the potential difference typically vary as a function of time as the capacitor charges or discharges. At any given instant, the magnitude of the potential difference across the capacitor is

$$V_{cap}(t) = \frac{q(t)}{C}, \tag{31.22}$$

where $q(t)$ is the magnitude of the charge on the plates at that instant. Because of this time-dependence, the current also depends on time. When solving Eq. 31.21 for circuits that contain capacitors, remember that the time-dependent current $I(t)$ and $q(t)$ are related by Eq. 27.2: $I(t) \equiv dq/dt$.

The Procedure box "Applying the loop rule in single-loop circuits" and **Table 31.2** summarize how to apply the loop rule to a single-loop circuit. Let us return to the circuit in Figure 31.34 and apply this procedure to obtain a relationship between the current and the emf. The figure already indicates a reference direction for the current (the arrowhead labeled I) and a direction of travel (the clockwise circular arrow). We shall begin our analysis at the bottom left corner. If we go clockwise, the first circuit element is the battery; because we are traveling from the negative to the positive terminal, the potential difference is $+\mathscr{E}$ (Figure 31.35*a* and Table 31.2). Next is the resistor; because we are traveling in the same direction as the

QUANTITATIVE TOOLS

reference direction for the current, the potential difference is $-IR$. Substituting these values in Eq. 31.21, we get

$$+\mathscr{E} - IR = 0 \tag{31.23}$$

or

$$I = \frac{\mathscr{E}}{R}. \tag{31.24}$$

PROCEDURE: Applying the loop rule in single-loop circuits

When applying the loop rule to a single-loop circuit consisting of resistors, batteries, and capacitors, we need to make several choices in order to calculate the current or the potential difference across each circuit element.

1. Choose a reference direction for the current in the loop. (This direction is arbitrary and may or may not be the direction of current, but don't worry, things sort themselves out in step 4.) Indicate your chosen reference direction by an arrowhead, and label the arrowhead with the symbol for the current (I).
2. Choose a direction of travel around the loop. This choice is arbitrary and separate from the choice of the reference direction for the current in step 1. (You may want to indicate the travel direction with a circular clockwise or counterclockwise arrow in the loop.)
3. Start traversing the loop in the direction chosen in step 2 from some arbitrary point on the loop. As you encounter circuit elements, each circuit element contributes a term to Eq. 31.21. Use Table 31.2 to determine the sign and value of each term. Add all terms to obtain the sum in Eq. 31.21. Make sure you traverse the loop completely.
4. Solve your expression for the desired quantity. If your solution indicates that $I < 0$, then the direction of current is opposite the reference direction you chose in step 1.

Table 31.2 Signs and values of potential differences across batteries and resistors (Figure 31.35)

Circuit element	Plus sign when traversing	Value
ideal battery	from − to +	$\mathscr{E}$
capacitor	from − to +	$q(t)/C$
resistor	opposite reference direction of current	IR



31.14 In the analysis of the circuit in Figure 31.34, we chose a clockwise reference direction for the current and a clockwise direction of travel. Redo the analysis using (*a*) a clockwise reference direction for the current and a counterclockwise direction of travel and (*b*) a counterclockwise reference direction for the current and a clockwise direction of travel.

EXERCISE 31.8 Series resistors

Consider the circuit shown in **Figure 31.36**, containing two resistors with resistances R_1 and R_2 and a battery with an emf $\mathscr{E}$. Determine the current in the circuit in terms of R_1, R_2, and $\mathscr{E}$.

Figure 31.36 Exercise 31.8.

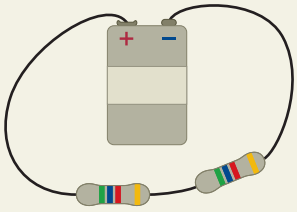
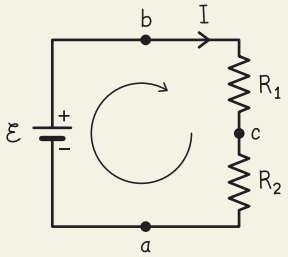


Figure 31.37



SOLUTION I begin by making a diagram, **Figure 31.37**, labeling the points a, b, and c, and choosing a clockwise reference direction for the current and a clockwise direction of travel. If I begin at point a, as I travel from the negative to the positive terminal, the potential difference is $+\mathscr{E}$. As I travel across resistor 1, I am traveling in the same direction as the reference direction for the current, so the potential difference is $-IR_1$. Next I travel across resistor 2 in the same direction as the reference direction for the current, so the potential difference is $-IR_2$. I now add all these terms and set the sum to zero (Eq. 31.21):

$$+\mathscr{E} - IR_1 - IR_2 = 0.$$

Solving this equation for the current yields

$$I = \frac{\mathscr{E}}{R_1 + R_2}. \checkmark$$

QUANTITATIVE TOOLS

Comparing the result of Exercise 31.8 with Eq. 31.11, we see that the effect on the current of placing two (or more) resistors in series is equivalent to adding the resistances, as we had concluded in Section 31.2. In other words, the same current would exist in the circuit if we replaced the two resistors in series with a single resistor having resistance

$$R_{eq} = R_1 + R_2. \tag{31.25}$$

By following the same line of reasoning, I can show that for more than two resistors in series, this result generalizes to

$$R_{eq} = R_1 + R_2 + R_3 + \cdots \text{ (resistors in series)}. \tag{31.26}$$

The resistance that could be used to replace a combination of circuit elements without altering the current from the battery is often referred to as the **equivalent resistance** of that part of the circuit.



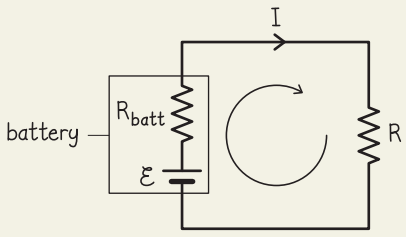
31.15 (*a*) In Figure 31.37, determine the potential difference between c and b by going counterclockwise from c to b. (*b*) For $\mathscr{E} = 9$ V, $R_1 = 10\ \Omega$, and $R_2 = 5\ \Omega$, calculate V_{cb} going counterclockwise from c to b.

EXAMPLE 31.9 Internal resistance

Even the best batteries dissipate some energy, which means that not all the chemical energy is converted to electric potential energy. We can take this dissipation into account by modeling a nonideal battery as consisting of an ideal battery of emf $\mathscr{E}$ in series with a small resistor R_{batt}, often called the *internal resistance* of the battery. The effect of this internal resistance is that the potential difference across the battery terminals is smaller than $\mathscr{E}$ when there is a current. When a nonideal battery is used in a circuit where the load is a resistor of resistance R, is the potential difference across the battery terminals greatest when the resistance of the load is high or when it is low?

❶ **GETTING STARTED** I begin by drawing a circuit diagram for such a nonideal battery connected to a load of resistance R (**Figure 31.38**). I need to determine an expression that shows me how the relative values of R and the internal resistance R_{batt} affect the potential difference across the battery terminals. I arbitrarily choose clockwise both for the reference direction for the current and for my travel around the circuit.

Figure 31.38



❷ **DEVISE PLAN** The potential difference across the nonideal battery terminals is the sum of the potential difference across the ideal battery, $\mathscr{E}$, and the potential difference across the internal resistance, $-IR_{batt}$. This latter term means I need to obtain a value for the current. I note that my circuit is equivalent to the two resistors in series in the circuit shown in Figure 31.37, and the final expression I obtained in Exercise 31.8 lets me express the current in terms of these two resistances. Thus I should be able to use this result to write an expression relating R and R_{batt} to V_{batt}.

❸ **EXECUTE PLAN** I use the solution I obtained in Exercise 31.8 to calculate the current in the circuit:

$$I = \frac{\mathscr{E}}{R_{batt} + R}.$$

I can now substitute this expression for I in the $-IR_{batt}$ term in my expression for V_{batt}:

$$V_{batt} = \mathscr{E} - IR_{batt} = \mathscr{E}\left(1 - \frac{R_{batt}}{R_{batt} + R}\right) = \mathscr{E}\left(\frac{R}{R_{batt} + R}\right).$$

When $R \gg R_{batt}$, the effect of R_{batt} is small and the potential difference across the battery terminals is essentially what it is in an ideal battery. When R is small, and especially when it is comparable to R_{batt}, then R_{batt} reduces the potential difference across the terminals. Thus the potential difference across the terminals of a nonideal battery is greatest when the resistance of the load is high. ✔

❹ **EVALUATE RESULT** If $R \gg R_{batt}$, the resistance in the circuit is great and hence the current is small. Because the current is small, the potential difference due to the internal resistance is small ($-IR_{batt}$) and hence the potential difference across the terminals does not differ significantly from its ideal value. This conclusion agrees with what my expression for V_{batt} shows.

QUANTITATIVE TOOLS

Electrical measuring instruments

Current, potential difference, and resistance can be measured with the following three electrical measuring devices. Often the three functions are combined into one instrument, called a *multimeter*.

Ammeter: To measure the current in a single-loop circuit (or in a branch of a multiloop circuit), an ammeter must be inserted into the loop or branch, as shown in **Figure 31.39*a*** or *c*. The ammeter then indicates the current passing through it.

Voltmeter: To measure the potential difference between two points, a voltmeter is connected to those two points (Figure 31.39*b*). Unlike current measurement, a potential difference measurement does not require breaking the circuit.

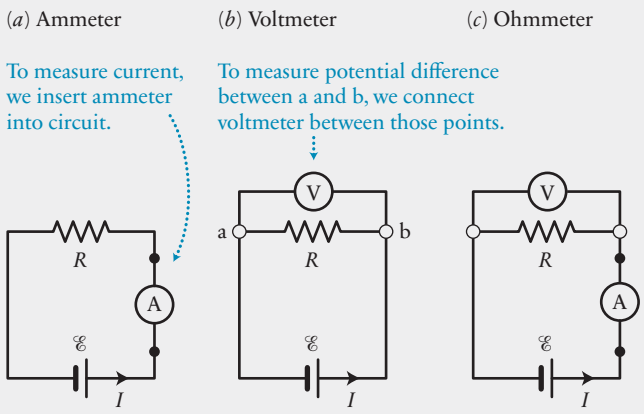
Ohmmeter: A voltmeter and an ammeter can be combined to measure a resistance, as shown in Figure 31.39*c*. In practice resistance is measured using a device called an *ohmmeter*. Such a device puts a known potential difference across the resistor to be measured, measures the resulting current though the resistor, and then uses Ohm's law to determine the resistance. To measure the resistance of a circuit element, that element must be disconnected from the circuit.

Voltmeters and ammeters typically affect the circuit under observation. For example, if the resistance R_A of the ammeter in Figure 31.39*a* is nonzero, the current in the circuit is affected by the insertion of the ammeter. The resistance of ammeters must therefore be very small compared to the other resistances in the circuit; an *ideal ammeter* has zero resistance.

Similarly, the resistance R_V of the voltmeter in Figure 31.39*b* must be very great compared to other resistances in the circuit to prevent charge carriers from flowing through the voltmeter and changing the current in the circuit; an *ideal voltmeter* has infinite resistance.

In either case, if we know the resistance of the device, it is possible to correct for its effect on the circuit.

Figure 31.39 Electrical measuring instruments.



According to the model of Example 31.9, an ideal battery is a battery with $R_{batt} = 0$. For a high-quality nonideal battery, R_{batt} can be extremely small—at most a few ohms. Thus, for ordinary loads, with resistances typically hundreds of ohms or more, the battery's internal resistance is not important.

A number of measuring instruments exist to measure electrical quantities in electric circuits and circuit elements. The instrument to measure currents is called an *ammeter*; the instrument to measure potential differences is called a *voltmeter*. By measuring both the current in a resistor and the potential difference across it, one can determine the resistance using Ohm's law. An *ohmmeter* is used to accomplish this task: Connecting a resistor between the terminals of an ohmmeter yields a reading for the resistance. See the box above on Electrical measuring instruments for more information on these devices.

31.7 Multiloop circuits

The analysis of multiloop circuits follows the same principles we laid out in the preceding section, but we need to consider what happens in multiple branches, junctions, and loops. The first thing to note is that the branch rule requires that in steady state, the current in any branch is the same everywhere along that branch. Therefore, in a circuit containing M branches, there are M distinct currents. When analyzing multiloop circuits, you should therefore always begin by identifying all the branches and labeling the current in each of those branches.

As we saw in Section 31.3, the **junction rule** states that in steady state the sum of the currents going into a junction must be equal to the sum of the currents going out of that junction:

$$I_{in} = I_{out} \quad \text{(steady state)}. \tag{31.27}$$

QUANTITATIVE TOOLS

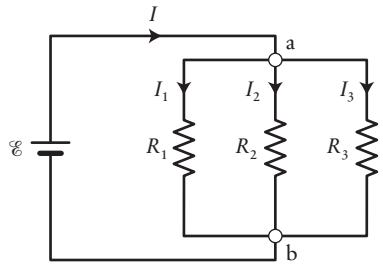


Figure 31.40 Circuit diagram for three resistors connected in parallel across an ideal battery.

As an example of a simple multiloop circuit, let us begin by considering three resistors connected in parallel across an ideal battery, as shown in Figure 31.40. Our goal will be to determine the equivalent resistance of the parallel combination of resistors. The circuit in Figure 31.40 has four branches, two junctions, and six loops (Can you identify them all?). Because there are four branches, we begin by labeling the currents in the four branches: the current in the battery, I, and the currents I_1, I_2, and I_3 through the three resistors. At junction a in Figure 31.40 we have $I_{\text{in}} = I$ and $I_{\text{out}} = I_1 + I_2 + I_3$, so the junction rule yields

$$I = I_1 + I_2 + I_3. \tag{31.28}$$

We can follow the same reasoning at junction b, but as you can easily verify, we would obtain exactly the same equation.

The potential difference V_{ba} is the same across each of the three resistors, so we can write for the currents in the resistors

$$I_1 = \frac{V_{\text{ba}}}{R_1}, \quad I_2 = \frac{V_{\text{ba}}}{R_2}, \quad I_3 = \frac{V_{\text{ba}}}{R_3}. \tag{31.29}$$

Substituting Eq. 31.29 into Eq. 31.28 gives

$$I = \frac{V_{\text{ba}}}{R_1} + \frac{V_{\text{ba}}}{R_2} + \frac{V_{\text{ba}}}{R_3} = V_{\text{ba}}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right). \tag{31.30}$$

If we replace the combination of the three parallel resistors by a single resistor having an equivalent resistance R_{eq}, we have

$$I = V_{\text{ba}}\left(\frac{1}{R_{\text{eq}}}\right). \tag{31.31}$$

The value of R_{eq} needed to obtain the same current I in the battery as with the three resistors connected in parallel is obtained by comparing Eqs. 31.30 and 31.31:

$$\frac{1}{R_{\text{eq}}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}, \tag{31.32}$$

which tells us that the reciprocal of the equivalent resistance of resistors connected in parallel equals the sum of the reciprocals of the individual resistances. Because the line of argument I used in deriving Eqs. 31.28–31.32 can be followed for any number of resistors, this statement is true for any number of resistors in parallel:

$$\frac{1}{R_{\text{eq}}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots \quad \text{(resistors in parallel)}. \tag{31.33}$$

For two resistors in parallel, the equivalent resistance can also be written as

$$R_{\text{eq}} = \frac{R_1 R_2}{R_1 + R_2} \quad \text{(two resistors in parallel)}. \tag{31.34}$$

Note that the equivalent resistance is smaller than either of the two individual resistances ($R_{\text{eq}} < R_1$ and $R_{\text{eq}} < R_2$). This is always true for resistors in parallel. As we discussed in Section 31.3, although it may seem paradoxical that combining multiple resistors reduces the equivalent resistance, combining resistors in parallel amounts to providing multiple paths for the current to follow. Viewed from this perspective, it is not surprising

QUANTITATIVE TOOLS

that increasing the number of branches in the circuit reduces the equivalent resistance of the circuit.

31.16 Let $\mathscr{E} = 9$ V, $R_1 = 3\ \Omega$, $R_2 = 10\ \Omega$, and $R_3 = 5\ \Omega$ in Figure 31.40. (*a*) What is the equivalent resistance of the three resistors? (*b*) What is the current in the battery?

The circuit in Figure 31.40 has more than one loop, but it is still a relatively simple circuit. To analyze more complex multiloop circuits, we need to derive a set of mathematical relationships among the unknown quantities (typically the currents in the various branches). Specifically, we need to have as many independent mathematical relationships as there are unknown quantities in the circuit.

One way to obtain a suitable set of equations is to apply the junction rule and the loop rule as many times as necessary to obtain a suitable number of equations. The junction rule and the loop rule are sometimes referred to as either *Kirchhoff's circuit rules* or *Kirchhoff's laws*, after the German physicist Gustav Kirchhoff (1824–1887). These rules do not contain any new physical principles, but instead are simply the application of principles we have already encountered—continuity (the junction rule) and Eq. 25.32 representing conservation of energy (the loop rule)—to an electric circuit in steady state.

Suppose we are interested in determining the currents in the circuit shown in **Figure 31.41**. Specifically, we want to determine the currents I_1, I_2, and I_3 in the three resistors. (The chosen reference directions in Figure 31.41 are arbitrary.) We wish to determine values for these currents in terms of the resistances R_1, R_2, and R_3 and the emfs of the batteries.

The circuit in Figure 31.41 has two junctions, three branches, and three loops. At junction b we have $I_{\text{in}} = I_1$ and $I_{\text{out}} = I_2 + I_3$, so, according to the junction rule,

$$I_1 = I_2 + I_3. \tag{31.35}$$

Applying the junction rule at f yields no new information because we would obtain the same equation. In general, it can be shown that in a circuit that contains N junctions, the junction rule yields $N - 1$ independent equations. This means that in a circuit with N junctions, we should apply the junction rule $N - 1$ times.

Next we need to use the loop rule. There are three unknown quantities in this problem (the three currents), so we need two equations in addition to Eq. 31.35 to determine the three currents. In other words, we need to apply the loop rule to two loops in the circuit.

First let us apply the loop rule to loop abdfea, choosing a clockwise travel direction around the loop:

$$V_{\text{ab}} + V_{\text{bd}} + V_{\text{df}} + V_{\text{fe}} + V_{\text{ea}} = 0. \tag{31.36}$$

Expressing the potential differences in this equation in terms of the emfs, resistances, and currents (see Table 31.2), we get

$$-I_1R_1 - \mathscr{E}_2 - I_2R_2 + 0 + \mathscr{E}_1 = 0. \tag{31.37}$$

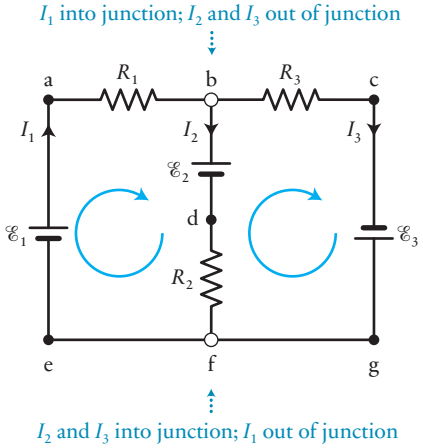
Next we apply the loop rule to loop bcgfdb, going clockwise around the loop:

$$V_{\text{bc}} + V_{\text{cg}} + V_{\text{gf}} + V_{\text{fd}} + V_{\text{db}} = 0, \tag{31.38}$$

or

$$-I_3R_3 + \mathscr{E}_3 + 0 + I_2R_2 + \mathscr{E}_2 = 0. \tag{31.39}$$

Figure 31.41 Using the junction and loop rules to determine the currents in the resistors in this multiloop circuit.



QUANTITATIVE TOOLS

PROCEDURE: Analyzing multiloop circuits

Here is a series of steps for calculating currents or potential differences in multiloop circuits.

1. Identify and label the junctions in the circuit.
2. Label the current in each branch of the circuit, arbitrarily assigning a direction to each current.
3. Apply the junction rule to all but one of the junctions. (The choice of which junctions to analyze is arbitrary; choose junctions that involve the quantities you are interested in calculating.)
4. Identify the loops in the circuit and apply the loop rule (see the Procedure box on page 1099) enough times to obtain a suitable number of simultaneous equations relating the unknowns in the problem. The choice of loops is arbitrary, but every branch must be in at least one of the loops. Traverse each loop in whichever direction you prefer, but be sure you traverse each loop completely and stick with the direction of travel and with the chosen directions of the currents.

There are several simplifications you can make during your analysis.

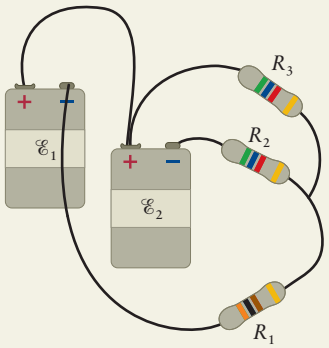
1. Multiloop circuits can sometimes be simplified by replacing parallel or series combinations of resistors by their equivalent resistances. If you can reduce the circuit to a single loop, you can solve for the current in the source. You may then need to "unsimplify" and undo the resistor simplification to calculate the current or potential difference across a particular resistor.
2. In general when solving problems, you should solve equations analytically before substituting known numerical values. When solving the simultaneous equations you obtain for multiloop circuits, however, you can often simplify the algebra if you substitute the known numerical values earlier on.

We now have three equations (Eqs. 31.35, 31.37, and 31.39) involving three unknowns (I_1, I_2, and I_3). (As stated earlier, the resistances and emfs are not unknowns. We consider R_1, R_2, R_3, $\mathscr{E}_1$, $\mathscr{E}_2$, and $\mathscr{E}_3$ to have certain values, although we are not given numerical values for them. Other problems might treat one or more of the resistances or emfs as unknowns, in which case some additional information, such as the values of some of the currents, would be needed.) Equations 31.35, 31.37, and 31.39 can be solved using standard algebraic techniques. Because every circuit is different, we shall not solve this particular problem. Complete solutions to such problems are provided in the examples below. All the physics of the circuit is expressed in Eqs. 31.35–31.39; the remainder of the solution simply involves algebra. The steps for analyzing multiloop circuits are summarized in the Procedure box above.

EXAMPLE 31.10 Multiloop circuit

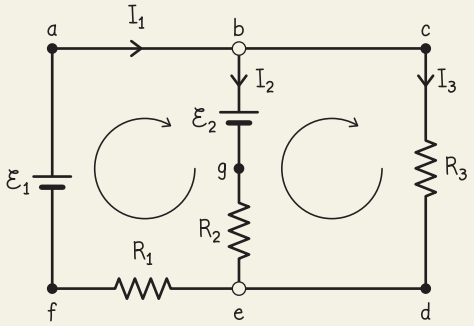
Consider the circuit shown in **Figure 31.42**. Determine the magnitude of the potential difference across R_1 if $\mathscr{E}_1 = \mathscr{E}_2 = 9.0$ V and $R_1 = R_2 = R_3 = 300\ \Omega$.

Figure 31.42 Example 31.10.



❶ **GETTING STARTED** I begin by drawing a circuit diagram for the circuit (**Figure 31.43**).

Figure 31.43



❷ **DEVISE PLAN** To solve this problem I follow the steps in the Procedure box on page 1104. There are two junctions in the circuit (b, e), three loops (abgefa, bcdegb, abcdefa), and three branches (efab, egb, edcb). I label the currents in the three branches I_1, I_2, and I_3. If I apply the junction rule to one of the junctions and the loop rule to two of the loops, I will obtain three equations that I can solve for the three unknowns I_1, I_2, and I_3. I choose loops abgefa and bcdegb and a clockwise travel direction in each of these loops. Once I know the current I_1, I can use Ohm's law (Eq. 31.11) to compute the potential difference across the resistor of resistance R_1. Because the values of the emfs and resistances are all the same, I set $\mathscr{E}_1 = \mathscr{E}_2 = \mathscr{E}$ and $R_1 = R_2 = R_3 = R$ to simplify the calculation.

❸ **EXECUTE PLAN** For junction b I have $I_{in} = I_1$ and $I_{out} = I_2 + I_3$, so applying the junction rule gives

$$I_1 = I_2 + I_3. \tag{1}$$

Applying the loop rule going clockwise around loop abgefa gives

$$0 - \mathscr{E} - I_2R - I_1R + \mathscr{E} = 0,$$

or

$$I_2 = -I_1. \tag{2}$$

Next I apply the loop rule going clockwise around loop bcdegb:

$$0 - I_3R + 0 + I_2R + \mathscr{E} = 0. \tag{3}$$

I now have three equations from which I can determine the three unknown currents.

Substituting Eq. 2 into Eq. 1 yields $I_1 = -I_1 + I_3$, or

$$I_3 = 2I_1. \tag{4}$$

Substituting this expression for I_3 and Eq. 2 into Eq. 3 yields

$$-2I_1R - I_1R + \mathscr{E} = 0.$$

Solving for I_1 then gives me

$$I_1 = \frac{\mathscr{E}}{3R}, \tag{5}$$

and substituting $\mathscr{E} = 9.0$ V and $R = 300\ \Omega$, I get

$$I_1 = \frac{9.0\text{ V}}{900\ \Omega} = 0.010\text{ A}.$$

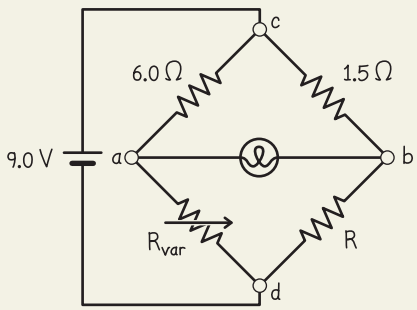
The magnitude of the potential difference across the resistor with resistance R_1 is therefore $|-I_1R_1| = (0.010\text{ A})(300\ \Omega) = 3.0$ V. ✔

❹ **EVALUATE RESULT** I can use my expression for I_1 to determine the potential difference between positions b and e. Going through branch efab, I have $V_{eb} = V_{ef} + V_{fa} + V_{ab} = -I_1R + \mathscr{E} + 0$. Substituting my expression for I_1 (Eq. 5), I get $V_{eb} = \frac{2}{3}\mathscr{E} = +6.0$ V. Going through branch egb, I have $V_{eb} = V_{eg} + V_{gb} = +I_2R + \mathscr{E}$. Because $I_2 = -I_1$ (Eq. 2), this expression gives the same result, $V_{eb} = +6.0$ V. Going through branch edcb and using Eq. 4, I get $V_{eb} = 0 + I_3R + 0 = +2I_1R$. Substituting Eq. 5, I obtain $V_{eb} = \frac{2}{3}\mathscr{E} = +6.0$ V. Because I obtain the same value for the potential difference each time, I am confident that my result for I_1 is correct.

EXAMPLE 31.11 Wheatstone bridge

The circuit shown in **Figure 31.44** includes a variable resistor, the resistance R_{var} of which can be adjusted, and a resistor of unknown value R. A circuit with such a network of resistors is called a *Wheatstone bridge* and can be used to determine the value of the unknown resistance R by adjusting the variable resistor so that the light bulb does not glow. The light bulb is initially glowing when R_{var} is set to 20 Ω, but it goes out when R_{var} is adjusted to 12 Ω. Determine the current in the battery when the light bulb is out.

Figure 31.44 Example 31.11.



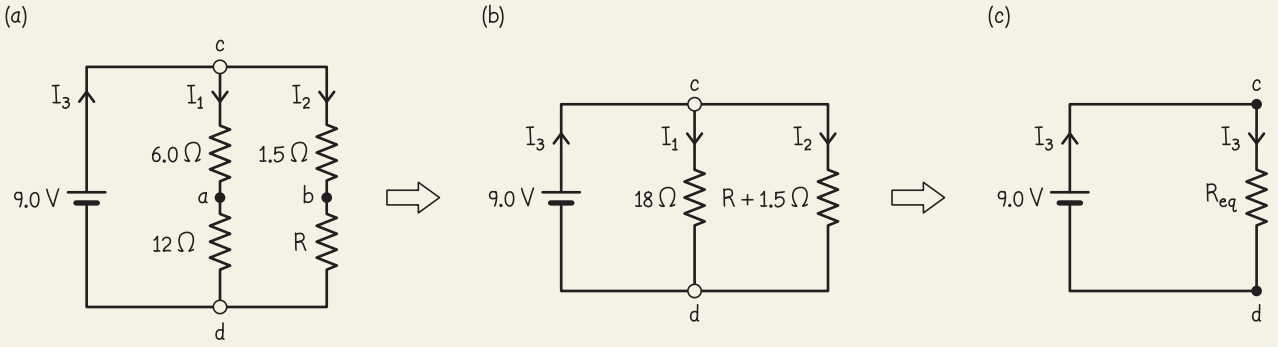
❶ **GETTING STARTED** The bulb glows when there is a current in it and stops glowing when the current in it is zero. In the latter case, it doesn't matter whether the bulb is connected or not, so I can omit it from the circuit and simplify the circuit diagram, as shown in **Figure 31.45*a*** on the next page. Because R_{var} is set to 12 Ω, I use that value rather than R_{var}. I use the label I_1 for the current in the branch containing the variable resistor, I_2 for the current in the branch containing R, and I_3 for the current in the battery, which is the current I need to determine.

❷ **DEVISE PLAN** In order for the current in the bulb to be zero, junctions a and b in Figure 31.44 must be at the same potential, and I must have $V_{ca} = V_{cb}$ and $V_{ad} = V_{bd}$. I can then use Ohm's law to express these two conditions in terms of the unknown resistance R and the currents I_1 and I_2, and solve the resulting set of equations for R. Once I know R, I can further simplify the circuit by replacing the sets of resistors connected in series in each branch by a single equivalent resistor. This yields the circuit shown in Figure 31.45*b*. To determine the current in the battery, I can then replace the two parallel resistors by a single equivalent resistor as in Figure 31.45*c*, and use Eq. 31.34 to determine its resistance.

(Continued)

QUANTITATIVE TOOLS

Figure 31.45



❸ **EXECUTE PLAN** Rearranging the Ohm's law equation (Eq. 31.11) to $V = IR$ and applying it to the two conditions $V_{ca} = V_{cb}$ and $V_{ad} = V_{bd}$, I have, when I travel clockwise in the circuit,

$$-(6.0\ \Omega)I_1 = -(1.5\ \Omega)I_2 \qquad (1)$$

$$-(12\ \Omega)I_1 = -RI_2. \qquad (2)$$

Solving Eq. 1 for I_1 gives me $I_1 = I_2/4$, and substituting this result into Eq. 2 and solving for R give me

$$R = 3.0\ \Omega.$$

With this value of R the resistance in the rightmost branch in Figure 31.45*b* becomes 4.5 Ω. The equivalent resistance of the two parallel resistors is, from Eq. 31.34,

$$R_{eq} = \frac{(18\ \Omega)(4.5\ \Omega)}{18\ \Omega + 4.5\ \Omega} = 3.6\ \Omega.$$

The current in the battery is then

$$I_3 = \frac{\mathscr{E}}{R_{eq}} = \frac{9.0\ \text{V}}{3.6\ \Omega} = 2.5\ \text{A}. \checkmark$$

❹ **EVALUATE RESULT** I can verify that $V_{ca} = V_{cb}$ by calculating I_1 and I_2 and substituting these values in Eq. 1. Applying the junction rule to junction c, I have $I_3 = I_1 + I_2$. I know that $I_1 = I_2/4$, so $I_2 = \frac{4}{5}I_3 = 2.0$ A and $I_1 = \frac{1}{5}I_3 = 0.5$ A. Using these values, I get $V_{ca} = -(6.0\ \Omega)(0.5\ \text{A}) = -3.0$ V and $V_{cb} = -(1.5\ \Omega)(2.0\ \text{A}) = -3.0$ V, so junctions a and b are at the same potential, as I expect when there is no current in the bulb.

Example 31.11 illustrates an alternative to our standard procedure for analyzing circuits. For circuits that contain many junctions and loops, it is sometimes easier to simplify the circuit by replacing part of it with a single equivalent resistance, as I did in Figure 31.45*b* of the example. When you do this, think about what question you ultimately want to answer about the circuit and about which parts of the circuit can be simplified without interfering with solving the problem. In Example 31.11, because what I wanted to determine was the current in the battery, I could replace everything outside the battery by a single equivalent resistor. However, if I had been asked to determine the currents at, say, junction b when the bulb was not glowing, I would have stopped simplifying the circuit at the stage shown in Figure 31.45*b*. Simplifying the circuit to just a single resistor and battery would have made it impossible to determine the current at b because the branch in which junction b is located would be gone from the circuit.

31.17 If R_{var} in Figure 31.44 is adjusted to a little less than 12 Ω, what is the direction of the current in the light bulb?

31.8 Power in electric circuits

At the beginning of this chapter, we gave an operational description of a single-loop circuit as converting electric potential energy from a power source to some other form of energy by driving a current in a load. We also found that the loop rule, which essentially embodies the idea of conservation of energy, is a powerful tool for analyzing circuits.

QUANTITATIVE TOOLS

Let us examine how rapidly the power source in **Figure 31.46** can deliver energy to its load. From Chapter 25 we know that the electrostatic work done on a charge carrier carrying charge q in a time interval Δt as it passes through the load from point a to point b is the negative of the potential difference across the load multiplied by q (Eq. 25.17):

$$W_q(\text{a} \rightarrow \text{b}) = -qV_{\text{ab}}. \tag{31.40}$$

Because V_{ab} is negative, the work done by the power source on the charge carrier is positive, increasing its energy. While it travels through the load, however, this energy is converted to other forms, and so the amount of energy converted in the load is given by

$$\Delta E = -qV_{\text{ab}}. \tag{31.41}$$

To determine the rate at which energy is converted, we substitute this expression into the expression for the average power (time rate of change of energy, Eq. 9.29). We are considering a steady-state situation, average and instantaneous power are the same, so we can drop the subscript av:

$$P \equiv \frac{\Delta E}{\Delta t} = \frac{-qV_{\text{ab}}}{\Delta t} = -IV_{\text{ab}}. \tag{31.42}$$

The form of energy to which the electric potential energy in a circuit is converted depends on the type of load. When the load is a light bulb, for example, the conversion is from electric potential energy to light and thermal energy. When the load is an electric motor, the conversion is to kinetic energy as, say, the blades of a fan start rotating. Here we're interested in circuits in which the load is a resistor, in which case the conversion is to thermal energy. Substituting $V_{\text{ab}} = -IR$ in Eq. 31.42 gives us, for the rate at which energy is dissipated in a resistor,

$$P = I^2R. \tag{31.43}$$

It can also be useful to know the rate at which energy is dissipated in terms of the potential difference across the resistor. Substituting $-V_{\text{ab}}/R$ for I in Eq. 31.43 yields

$$P = \frac{V_{\text{ab}}^2}{\text{R}}. \tag{31.44}$$

Keep in mind that although Eq. 31.42 is valid for any electrical device, we have derived Eqs. 31.43 and 31.44 using Ohm's law for resistors, in which electric potential energy is converted to thermal energy.

We can apply a similar reasoning to the power source. Let's consider the case of a battery as the power source. When a charge carrier carrying charge q moves from the negative terminal to the positive terminal inside the battery, its electric potential energy increases at a rate given by

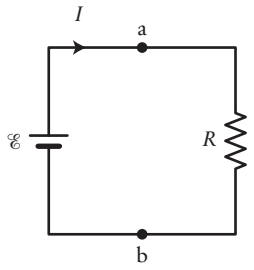
$$P = \frac{q\mathscr{E}}{\Delta t} = I\mathscr{E}. \tag{31.45}$$

For an ideal battery $\mathscr{E} = IR$ (Eq. 31.24), and so Eq. 31.45 becomes

$$P = I\mathscr{E} = I^2R. \tag{31.46}$$

Comparing Eqs. 31.46 and 31.43, we see that the rate at which electric potential energy increases in the battery is equal to the rate at which

Figure 31.46 A battery delivers energy to a load connected to it.



electric potential energy is dissipated in the load. For a nonideal battery, some energy is dissipated inside the battery. As we have seen in Example 31.9, we can account for this dissipation by attributing an internal resistance R_{batt} to the battery. This internal resistance decreases the potential difference across the terminals: $V_{ba} = \mathscr{E} - IR_{batt}$, so $\mathscr{E} = V_{ba} + IR_{batt}$. Substituting this expression and $V_{ba} = IR$ into Eq. 31.45 yields

$$P = I\mathscr{E} = IV_{ba} + I^2R_{batt} = I^2R + I^2R_{batt}. \tag{31.47}$$

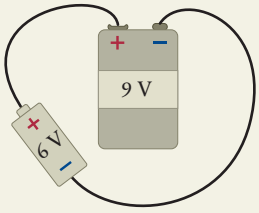
So for a nonideal battery the rate at which chemical energy is converted is equal to the sum of the rates at which energy is dissipated in the load and inside the battery, as we would expect.

31.18 The SI units of power suggested by Eqs. 31.42 and 31.43 are A · V and $A^2 \cdot \Omega$, respectively. Show that these SI units are equivalent to the derived SI unit for power, the watt.

EXAMPLE 31.12 Battery to battery

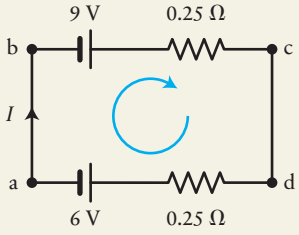
A 9.0-V and a 6.0-V battery are connected to each other (Figure 31.47). Each battery has an internal resistance of 0.25 Ω. At what rate is energy dissipated in the 6.0-V battery?

Figure 31.47 Example 31.12.



❶ **GETTING STARTED** I begin by drawing a circuit diagram for the two-battery combination, representing the internal resistance by two resistors in series with the batteries (Figure 31.48). I note that the negative terminals of the batteries are connected to each other, and the positive terminals are connected to each other. I arbitrarily choose clockwise both for the reference direction for the current and for my travel around the circuit.

Figure 31.48



❷ **DEVISE PLAN** Because this circuit contains only one loop and no junctions, the current is the same everywhere in the circuit. To determine the current, I apply the loop rule and solve the resulting equation for the current. I can then apply Eq. 31.43 to determine the rate at which energy is dissipated in the 6.0-V battery.

❸ **EXECUTE PLAN** As I travel clockwise around the circuit starting at a, the loop rule yields

$$9.0\ \text{V} - I(0.25\ \Omega) - I(0.25\ \Omega) - 6.0\ \text{V} = 0$$

$$I = \frac{9.0\ \text{V} - 6.0\ \text{V}}{0.50\ \Omega} = \frac{3.0\ \text{V}}{0.50\ \Omega} = 6.0\ \text{A}.$$

The fact that I get a positive value for the current tells me that my assumed reference direction for the current (clockwise) is correct. Energy is dissipated in the 6.0-V battery as resistive losses in the internal resistance. Using Eq. 31.43 for the rate at which energy is dissipated in a resistor, I get

$$P = I^2R = (6.0\ \text{A})^2(0.25\ \Omega) = 9.0\ \text{W}. \checkmark$$

❹ **EVALUATE RESULT** Even though the internal resistance is small, this power is substantial because the current is large. The current is large because there is very little resistance in the circuit. It's not a good idea to connect two batteries of different emf values in this manner!

31.19 (*a*) In Example 31.12, how would the answer change if we had chosen a counterclockwise travel direction around the circuit? (*b*) At what rate is energy dissipated in the 9-V battery?

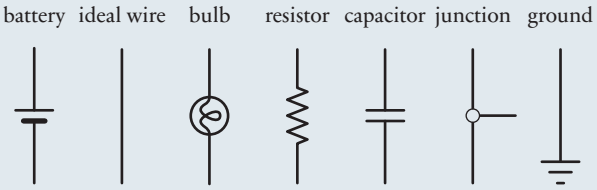
QUANTITATIVE TOOLS

Chapter Summary

The basic circuit (Section 31.1)

Concepts An **electric circuit** is an interconnection of electrical components (*circuit elements*). A **loop** is any closed conducting path through a circuit. A **power source** is any device that provides electric potential energy to an electric circuit. The potential difference across the terminals of the power source drives charge carriers through the circuit and thereby creates a current in the circuit. The **load** in an electric circuit is all the circuit elements connected to the power source. In the load, electric potential energy is converted to other forms of energy.

Quantitative Tools Standard representations of some common circuit elements used in circuit diagrams:



Current and resistance (Sections 31.2, 31.4, 31.5, 31.8)

Concepts A circuit is in **steady state** when the current has a constant value at all points in the circuit. The **current continuity principle** states that in steady state, the current is the same at all locations along a single-loop circuit.

The **resistance** of a circuit element is a measure of the potential difference across that element for a given current in it. The SI unit of resistance is the **ohm** Ω, where $1\Omega \equiv 1\ \text{V/A}$. The **conductivity** σ of a material is a measure of its ability to conduct charge carriers through the material and is measured in SI units of $\text{A}/(\text{V}\cdot\text{m})$.

When there is a current in a conductor, there must be an electric field in the conductor to cause the current. In a conductor of uniform cross section at steady state, this electric field has the same magnitude everywhere and is parallel to the walls of the conductor.

The **drift velocity** $\vec{v}_d$ of the charge carriers in a metal in which a current is present is the average velocity with which the carriers are moving.

The **current density** $\vec{J}$ in an electrical conductor is a vector whose direction is the same as the drift velocity direction when the charge carriers are positive and opposite the drift velocity direction when the charge carriers are negative.

Quantitative Tools The magnitude J of the **current density** is

$$J \equiv \frac{|I|}{A} = n|q|v_d, \tag{31.5}$$

where n is the number of charge carriers per unit volume of the conductor and v_d is the magnitude of the drift velocity of the charge carriers. If the electric field in a metal is $\vec{E}$, the **drift velocity** of the charge carriers (free electrons) is

$$\vec{v}_d = -\frac{e\vec{E}}{m_e}\tau, \tag{31.3}$$

where m_e is the mass of an electron and τ is the average time interval between collisions of the electrons with the metal ions.

The **conductivity** of a metal is

$$\sigma \equiv \frac{J}{E} = \frac{ne^2\tau}{m_e}. \tag{31.8, 31.9}$$

The **resistance** of any circuit element is

$$R \equiv \frac{V}{I}, \tag{31.10}$$

where V is the applied potential difference across the element and I is the resulting current in the element. If the curve of I versus V for a conductor is a straight line, the conductor obeys **Ohm's law**, $I = V/R$ (Eq. 31.11), and we call it *ohmic*.

The resistance of a conductor of length ℓ and cross-sectional area A is

$$R = \frac{\ell}{\sigma A}. \tag{31.14}$$

The rate at which energy is dissipated in a resistor is

$$P = I^2R. \tag{31.43}$$

The rate at which a power source delivers energy is

$$P = I\mathscr{E}, \tag{31.45}$$

where $\mathscr{E}$ is the emf of the source.

Circuit loops (Sections 31.3, 31.6, 31.7)

Concepts A **junction** is a location in a circuit where more than two wires are connected. A **branch** in a circuit is a conducting path between two junctions that is not intercepted by another junction. According to the **branch rule,** the current in a branch of a multiloop circuit is the same throughout the branch.

Two or more circuit elements are connected in *series* if there is only a single current path through them and the charge carriers flow through one element after the other. The potential difference across circuit elements connected in series is equal to the sum of the individual potential differences across each circuit element.

Two or more circuit elements are connected in *parallel* if their ends are connected to the same two junctions. The potential differences across circuit elements connected in parallel are always equal.

Quantitative Tools By the **loop rule,** the sum of the emfs and the potential differences around any loop under steady-state conditions is zero:

$$\Sigma\mathscr{E} + \Sigma V = 0. \tag{31.21}$$

For a steady-state multiloop circuit, the **junction rule** states that the sum of the charge carriers going into a junction equals the sum of the carriers coming out of that junction:

$$I_{\text{in}} = I_{\text{out}}. \tag{31.27}$$

See the Procedure boxes "Applying the loop rule in single-loop circuits" and "Analyzing multiloop circuits."

The **equivalent resistance** of resistors *in series* is given by

$$R_{\text{eq}} = R_1 + R_2 + R_3 + \cdots. \tag{31.26}$$

The **equivalent resistance** of resistors *in parallel* is given by

$$\frac{1}{R_{\text{eq}}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots. \tag{31.33}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

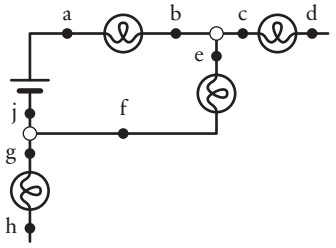
Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

Unless instructed otherwise, assume all batteries and all wires in circuits are ideal and thus have zero (internal) resistance.

31.1 The basic circuit

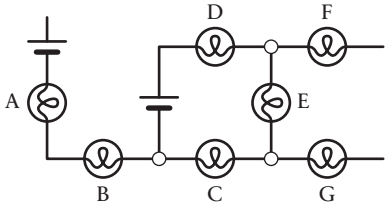
1. A battery that initially contains 3.0×10^{24} electrons is used to power a light bulb for some time interval, pumping 1.1×10^{24} electrons through the bulb during that interval. How many electrons remain in the battery? •
2. How can the energy conversions in an electric circuit be represented in terms of two generalized circuit elements? •
3. A battery and four identical light bulbs are arranged in the circuit of Figure P31.3. Rank the current magnitudes at the nine lettered locations from greatest to smallest. •

Figure P31.3



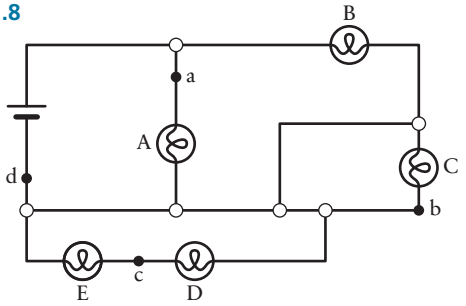
4. Which of the seven identical light bulbs in the circuit of Figure P31.4 light up? •

Figure P31.4



5. You have a light bulb, a battery, and one wire (which you cannot cut into two pieces). Draw the four ways of connecting these elements so that the bulb lights up. ••
6. Name any devices other than the three mentioned in Section 31.1—battery, solar cell, and electric generator—that can act as a power source in an electric circuit. ••
7. Draw a circuit diagram for a typical home hair dryer. To which form (or forms) of energy is electric potential energy converted when you use the dryer? ••
8. Figure P31.8 shows five identical light bulbs A–E connected to a battery. Initially, some of the bulbs light up

Figure P31.8



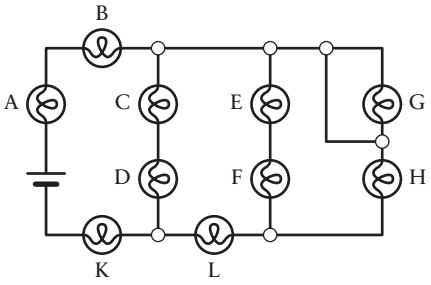
because the two terminals of each bulb are connected to opposite terminals of the battery. If any wire in the circuit is cut, some bulbs may go out. Which bulbs, if any, go out when a single cut is made at (*a*) location a, (*b*) location b, (*c*) location c, and (*d*) location d? •••

9. You and a colleague want to light a light bulb using a parallel-plate capacitor as a power source. Your colleague objects to this design, stating that the bulb cannot light up if the circuit is not continuous and the circuit is not continuous because the air gap between the capacitor plates constitutes a break in the circuit. He argues that charge carriers do not jump across the gap, so how can there be any current? What is your counterargument? •••

31.2 Current and resistance

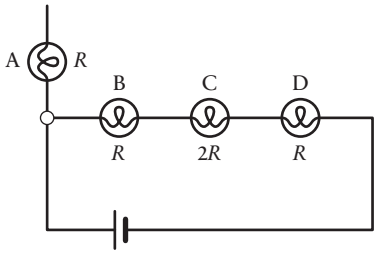
10. Which of the light bulbs (if any) in Figure P31.10 are connected in series to each other? •

Figure P31.10



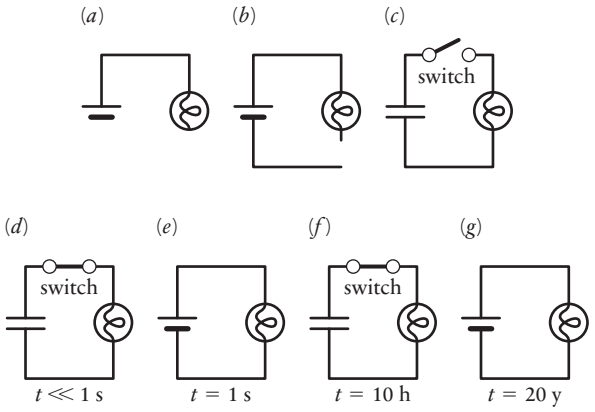
11. (*a*) In the circuit of Figure P31.11, do all four bulbs light up? (*b*) Rank the bulbs according to how brightly they light up, brightest first. •

Figure P31.11



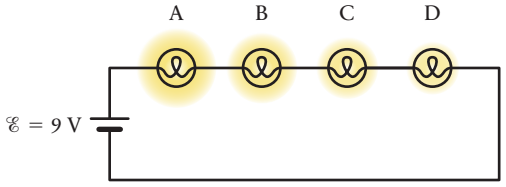
12. All the power sources (batteries or capacitors) in Figure P31.12 start with a potential difference of 9.0 V between their terminals, and all of them run down (decrease in potential difference) over time. Rank the brightness of these identical light bulbs from least bright to most bright. The *t* values in parts *d* through *g* refer to the time interval that has passed since the power source was connected to the load. ••

Figure P31.12



13. In Figure P31.13, bulb A is brighter than bulb B, B is brighter than C, and C is brighter than D. Is it possible for the potential difference across bulb C to be (*a*) 3 V and (*b*) 2 V? ••

Figure P31.13



14. When light bulbs A and B are connected in series to a battery, bulb A glows brightly and bulb B glows dimly. You remove bulb B so that the circuit contains just the battery and bulb A, and you connect bulb B to an identical battery. With the two bulbs connected to identical batteries but in separate circuits, do the bulbs glow equally brightly now? ••

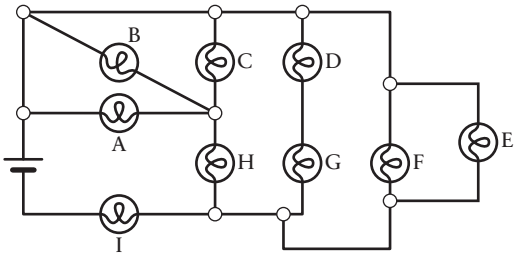
15. The bulbs A, B, C, and D are connected in series with a 12.0-V battery, and the steady current through bulb A is 1.0 A. Bulbs B and C, respectively, produce twice and three times as much light and thermal energy as bulb A and bulb D produces the same as bulb B. In 2.0 s, how much energy is dissipated as light and thermal energy by each bulb? •••

31.3 Junctions and multiple loops

16. Wires A, B, C, D, and E meet at a junction point O. The current in wire D and E is twice as in wire A and four times as in C, respectively. The current in wire A is 4.2 mA into the junction, and the current in wire C is 5.6 mA out of the junction. What is the current in wire B, and in what direction—into or out of the junction—is the current? •

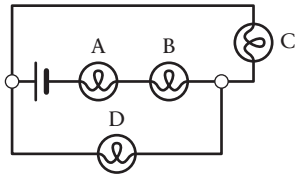
17. Which (if any) of the light bulbs in Figure P31.17 are connected to other bulbs in parallel? •

Figure P31.17



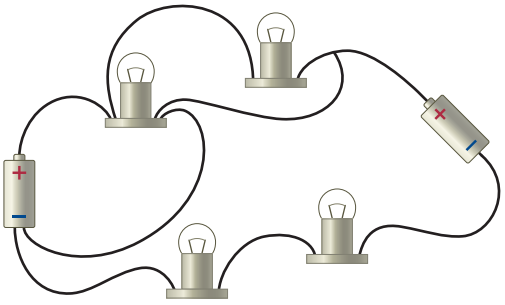
18. Which of the light bulbs in Figure P31.18 are connected to each other in series? Which are connected to each other in parallel? •

Figure P31.18



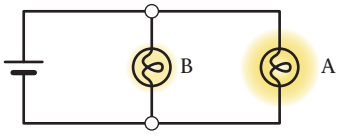
19. Draw a circuit diagram for the circuit of Figure P31.19. ••

Figure P31.19



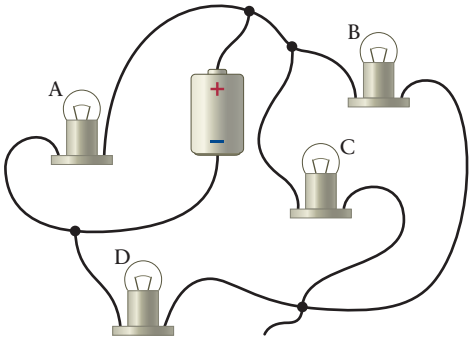
20. In the circuit of Figure P31.20, bulb A is bright and bulb B is dim. Which bulb (*a*) has the greater potential difference across it, (*b*) carries the greater current, and (*c*) has the greater resistance? ••

Figure P31.20



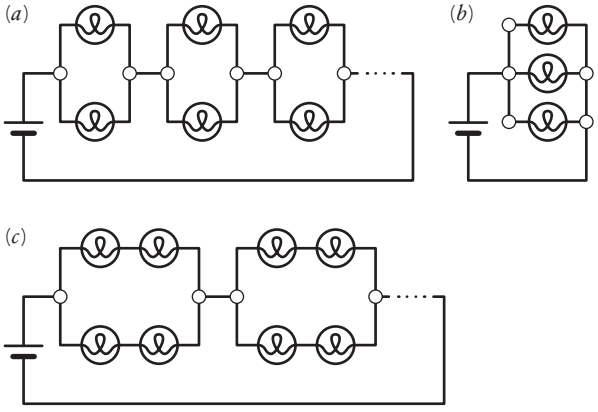
21. (*a*) Draw a circuit diagram for the circuit shown in Figure P31.21, which consists of a battery and four identical light bulbs. (*b*) Do all the bulbs light up? (*c*) Which bulb is brightest? Dimmest? (*d*) Which bulbs are connected to each other in parallel? Which are connected in series? ••

Figure P31.21



22. Decades ago, holiday lights were wired in series, so that if one bulb in a string burned out, all the lights in the string went dark because the burned-out bulb interrupted the circuit. Today's lights are wired at least partially in parallel, so that if one light goes out, many others in the string remain lit. In which circuits in Figure P31.22 does one bulb burning out leave all other bulbs lit? In which does one bulb burning out cause one or more of the other bulbs to go out? ••

Figure P31.22

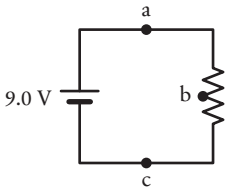


23. Are the lights in a house wired in series, in parallel, or a combination of the two? •••

31.4 Electric fields in conductors

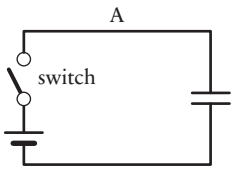
24. In Figure P31.24, the wire connecting the positive battery terminal to the top end of the resistor is 100 mm long, and the wire connecting the negative battery terminal to the bottom end of the resistor is also 100 mm long. If the resistor is 10 mm long, what are the electric field magnitudes at locations a, b, and c? •

Figure P31.24



25. Once the switch in the circuit in Figure P31.25 is closed, the circuit can be used to charge the capacitor. Sketch the electric field inside wire A (*a*) when the switch is open, (*b*) at the instant just after the switch is closed, and (*c*) at some instant long after the switch is closed. •

Figure P31.25



26. (*a*) Given a copper wire with radius 850 μm carrying a uniform current of 25 mA, calculate the magnitude of electric field applied to this wire. (*b*) We have given two types of wire made up of copper and aluminum. If these wires are of the same length, radius, and carry same current, which wire has higher electric field? ••

27. You want to use a length of insulated rigid metal rod to discharge the plates of a parallel-plate capacitor without allowing the electric field in the rod to exceed 1000 N/C. The area A of each square plate is 1.00 m^2, the plate separation distance is $d = 100$ mm, and there is no dielectric material between the plates. Initially, each plate holds a charge of magnitude 4.51×10^{-8} C. If one end of the rod is attached to the center of one edge of plate 1, describe all the positions on plate 2 to which the other end of the rod could be attached. Illustrate your answer with a sketch. •••

31.5 Resistance and Ohm's law

28. Given a copper wire of length 100 cm, radius 0.8 mm, and resistivity $1.68 \times 10^{-8}\ \Omega \cdot$m, find its resistance and current through the wire if connected to a 20.0-V battery. If the wire is stretched uniformly to 4 times its original length, what will be its new resistance? •

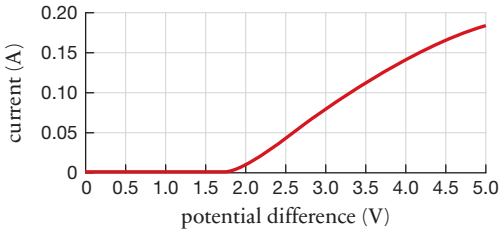
29. To connect two devices, you need a 40 m long copper wire. If the needed resistance of the wire is 0.5 Ω, then what should be the diameter of the wire? If the voltage drop across the wire is 10.0-V, then what is the current flow to each device? [$\rho_{\text{copper}} = 1.68 \times 10^{-8}\ \Omega \cdot$m] •

30. Determine the current density when 5.0 A of current is flowing through a copper wire of radius 0.6 mm. By what amount does current density change if the wire's radius is doubled?

31. How strong must an electric field in a metal be in order for electrons in the field to have a drift speed of 10 mm/s if the time interval between electron-ion collisions is 1.0×10^{-14} s? •

32. Figure P31.32 is a graph of the current through a light-emitting diode as a function of the potential difference across the diode. What is the resistance of the diode at a potential difference of 3.0 V? ••

Figure P31.32



33. *Gauge* is a term used to describe the size of a wire: The greater the gauge, the smaller the wire diameter. At room temperature, 6-gauge wire has a diameter of 4.115 mm, and 22-gauge wire has a diameter of 0.644 mm. For copper, how long would a wire of each gauge need to be to have a resistance of 0.5 Ω [$\sigma_{\text{copper}} = 5.9 \times 10^7\ \text{A}/(\text{V} \cdot \text{m})$]? ••

34. Two wires are made of the same material. If both are at the same temperature, but one has twice the diameter and three times the length of the other, which has the greater resistance, and by what factor? ••

35. When a potential difference of 1.00 V is maintained between opposing faces of a metal that has a charge carrier number density of $n = 6.60 \times 10^{28}$ charge carriers/m^3, the cube carries current $I = 6.10 \times 10^5$ A. What is the average time interval τ between collisions in this metal if each side of the cube is 10.0 mm long? ••

36. What is the magnitude of the applied electric field inside an aluminum wire of radius 1.0 mm that carries a 4.0-A current [$\sigma_{\text{aluminum}} = 3.6 \times 10^7\ \text{A}/(\text{V} \cdot \text{m})$]? ••

37. For each of these changes in a metal, predict whether the average time interval between collisions of an electron with a lattice ion increases or decreases: (*a*) spatial density of the lattice ions is increased, (*b*) size of the lattice ions is decreased, (*c*) charge of the lattice ions is increased. ••

38. In a copper wire that has a diameter of 1.63 mm, the drift velocity is 7.08×10^{-4} m/s. If we assume one free electron per copper atom, what are (*a*) the current in the wire and (*b*) the current density? (*c*) This wire connects a light bulb, power source, and switch in a circuit, and the distance between the switch and bulb is 3.00 m. How long does it take for an electron initially at the switch to reach the bulb? ••

39. In a particle accelerator, the particles in a beam of protons are traveling toward a target at a speed equal to $0.100c_0$. If the beam has a radius of 0.100 μm and carries a current of 2.00 nA, how many protons strike the target in 1.00 s? Assume the target is large enough to intercept the entire beam. •

40. When you step on the brake pedal in your car, charge carriers flow from the battery to the rear brake lights. Suppose the wire connecting the switch at the pedal to the brake lights is made of copper and has a diameter of 1.1 mm. If the current through the wire is 2.0 A, how long on average does it take an electron to travel from the switch at the pedal to one of the brake lights? The number density of free electrons in copper is $n = 8.4 \times 10^{28}$ electrons/m^3. ••

41. Even though silver is a better electrical conductor than copper, most electrical cables are made from copper. The main reason is cost: The per-kilogram price of silver is about 100 times the per-kilogram price of copper. If you want a length ℓ of wire to have a resistance R, by what factor does the cost of silver exceed that of copper? (The mass density of silver is $\rho_S = 10{,}490$ kg/m^3, and that of copper is $\rho_C = 8960$ kg/m^3.) ••

42. You have a piece of copper wire and a piece of carbon rod, each 1.5 m long and each having a cross-sectional area of 8.0×10^{-6} m^2. When you connect the copper wire to the terminals of a 9.0-V battery, the current in the wire is I. If you want to connect the carbon rod to another battery and

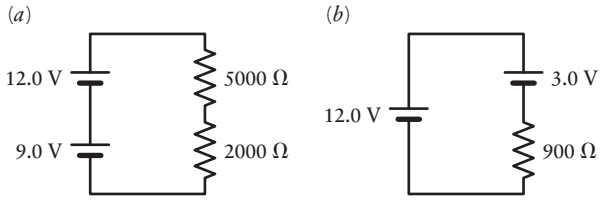
have the same current I in the wire, what must the potential difference of the battery be [$\sigma_{copper} = 5.9 \times 10^7$ A/(V·m), $\sigma_{carbon} = 7.3 \times 10^4$ A/(V·m)]? ●●

43. An electric field of magnitude 4.50×10^2 V/m is created in a wire that is 300 mm long and has a radius of 1.00 mm. The number density of the charge carriers in the wire is 1.20×10^{27} charge carriers/m^3. How much energy is gained by the lattice of metal ions in the time interval it takes for the average electron to travel the length of the wire? ●●●

31.6 Single-loop circuits

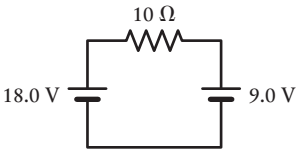
44. What are the magnitude and direction of the current in each circuit in Figure P31.44? ●

Figure P31.44



45. What are the magnitude and direction of the current in the circuit in Figure P31.45? ●

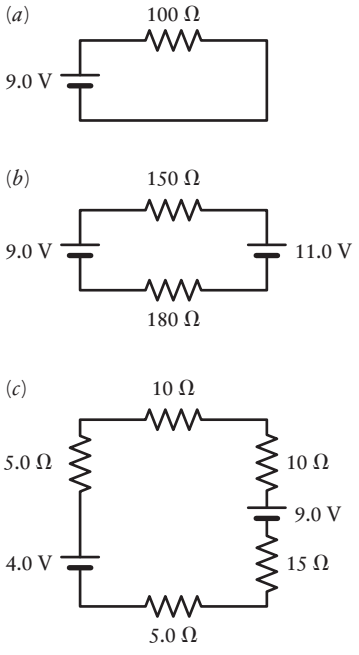
Figure P31.45



46. Three resistors are connected in series to a battery providing an emf of 24 V. If the resistances are $R_1 = 8\ \Omega$, $R_2 = 5\ \Omega$, and $R_3 = 3\ \Omega$, what is the current that flows through the 8-Ω resistor (*a*) if the battery is ideal and (*b*) if the battery has an internal resistance of 2.0 Ω? ●

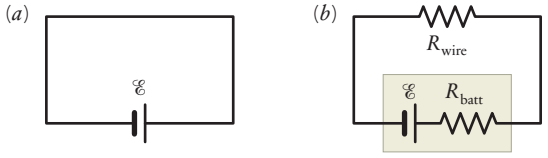
47. What are the magnitude and direction of the current in each circuit in Figure P31.47? ●●

Figure P31.47



48. When using the loop rule, what problems do you encounter with the circuit in Figure P31.48*a*? What do you expect for the current in the more realistic circuit in Figure P31.48*b*, which includes a small nonzero resistance in the wire and internal resistance in the battery? ●●

Figure P31.48

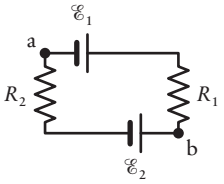


49. A nonideal battery connected to a resistor R_1 in a circuit results in the current I_1, if the same battery connected to a resistor R_2 gives current I_2. (*a*) What should be the value of R_2 for which the internal resistance is zero? (*b*) If $R_2 = 2R_1$ and $I_1 = 2I_2$, calculate the emf of the battery in terms of R_1 and I_1. ●●

50. List all the different resistances one can obtain using three resistors, each of value 50 Ω. ●●

51. The potential difference between positions a and b in Figure P31.51 is 6 V, and $R_1 = 4.0\ \Omega$, $\mathscr{E}_1 = 10.0$V, and $\mathscr{E}_2 = 5.0$ V. What is the value of the resistance R_2? ●●

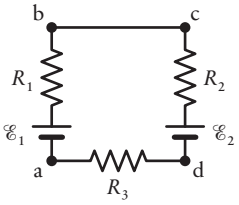
Figure P31.51



52. A typical car battery can be modeled as an ideal source $\mathscr{E}$ connected in series with an internal resistance R_{batt}. Using a good battery of this design to "jump start" a dead one can be dangerous if done incorrectly. (*a*) If you want to connect a good battery to a dead one in such a way that you have the smallest possible current through the dead one, should you connect them in series (+ to − and − to +) or in parallel (+ to + and − to −)? Assume you have a good battery for which $\mathscr{E}_{good} = 12.0$ V and $R_{batt,\,good} = 0.0200\ \Omega$ and a dead one for which $\mathscr{E}_{dead} = 11.0$ V and $R_{batt,\,dead} = 0.200\ \Omega$. Calculate the current through the batteries connected (*b*) in series and (*c*) in parallel. ●●

53. For the circuit shown in Figure P31.53, calculate (*a*) the magnitude and direction of the current and (*b*) the potential differences V_{ab}, V_{bc}, V_{cd}, and V_{da}. Use these values: $R_1 = 4.0\ \Omega$, $R_2 = 2.0\ \Omega$, $R_3 = 6.0\ \Omega$, $\mathscr{E}_1 = 4$ V, and $\mathscr{E}_2 = 8$ V. ●●

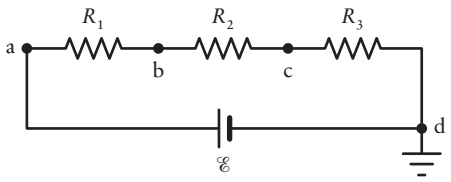
Figure P31.53



54. The internal resistance of a battery is relatively small when the battery is new but increases as the battery ages. When a new 12.0-V battery is attached to a 100-Ω load, the potential difference across the load is 11.9 V. After the circuit has operated for a while, the potential difference across the load is 11.5 V. By how much has the internal resistance of the battery changed? ••

55. (*a*) What is the equivalent resistance of the circuit in Figure P31.55? Use the values $R_1 = 10\ \Omega$, $R_2 = 60\ \Omega$, and $R_3 = 5\ \Omega$. (*b*) What is the current in the circuit? Assume $\mathscr{E} = 15$ V. (*c*) The electric potential at location d is defined to be 0, because the negative terminal of the battery is attached to ground. What is the electric potential at locations *a*, *b*, and *c*? ••

Figure P31.55

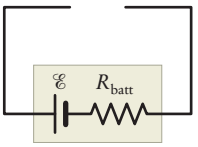


56. Resistors 1 and 2—$R_1 = 40\ \Omega$, $R_2 = 70\ \Omega$—are connected in series to a 4.5-V battery. (*a*) What is the potential difference across resistor 1? (*b*) If you decrease the value of R_1, what happens to the current in the circuit and to the potential difference across resistor 1? ••

57. A light bulb has resistance $R_{\text{bulb}} = 5.0\ \Omega$ and should be operated at a potential difference of $V_{\text{bulb}} = 3.0$ V. If you must use this bulb in a circuit powered by a battery of emf $\mathscr{E} = 9.0$ V, how much resistance must you add in series to the circuit in order to have a 3.0-V potential difference across the bulb? ••

58. You must complete the circuit of Figure P31.58 in such a way that it draws a current of 0.300 A from the battery. The battery maintains a potential difference of 10.0 V with no load, but has an internal resistance of $R_{\text{batt}} = 18.0\ \Omega$. The only material you have is 20.0 mm^3 of nichrome, and you must use all of it. If your plan is to form the nichrome into a cylindrical resistor, what must the cylinder length and cross-sectional area be? [The conductivity of nichrome is $6.7 \times 10^5\ \text{A/(V·m)}$.] •••

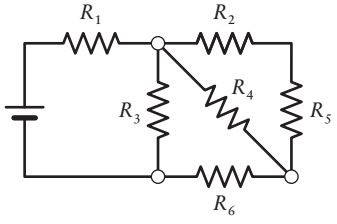
Figure P31.58



31.7 Multiloop circuits

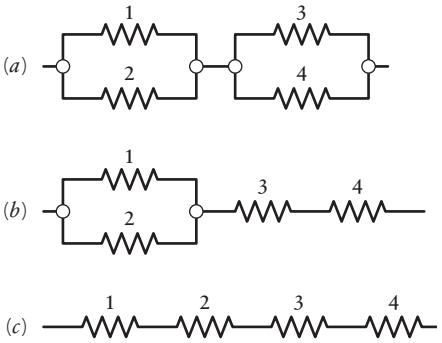
59. What is the equivalent resistance of the circuit in Figure P31.59? Use the values $R_1 = 4.0\ \Omega$, $R_2 = 2.0\ \Omega$, $R_3 = 1.0\ \Omega$, $R_4 = 3\ \Omega$, $R_5 = 1.0\ \Omega$, and $R_6 = 2.5\ \Omega$. •

Figure P31.59



60. Figure P31.60 shows three circuits containing four identical resistors, each having resistance *R*. Which circuit has the smallest equivalent resistance? Which has the greatest equivalent resistance? •

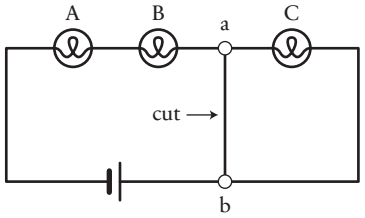
Figure P31.60



61. A nonideal ammeter that has an internal resistance of 0.503 Ω is connected in series with a 3.00-V battery and a 40.0-Ω resistor. By what percentage does the measured current differ from that of the same circuit with no ammeter? •

62. In Figure P31.62, the brightness of each bulb depends on the magnitude of the current through it. Rank these identical bulbs according to brightness, brightest first, (*a*) before the wire ab connecting the two junctions is cut and (*b*) after the wire is cut. (*c*) Does cutting the wire increase or decrease the brightness of A? Of B? Of C? ••

Figure P31.62

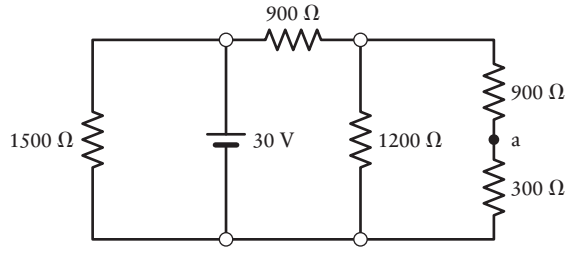


63. Using only 10.0-Ω resistors (but as many as you like), build a circuit that has a resistance of 27.5 Ω. ••

64. A copper wire has a diameter of 0.20 mm and is $\ell_{\text{wire}} = 10$ m long. (*a*) What is the resistance of the wire? (*b*) The wire is cut into *N* identical pieces, and the pieces are connected in parallel to form a single resistor. What is the minimum value of *N* that gives this resistor a resistance smaller than $R_{\text{resistor}} = 1.0\ \Omega$? [The conductivity of copper is $5.9 \times 10^7\ \text{A/(V·m)}$.] ••

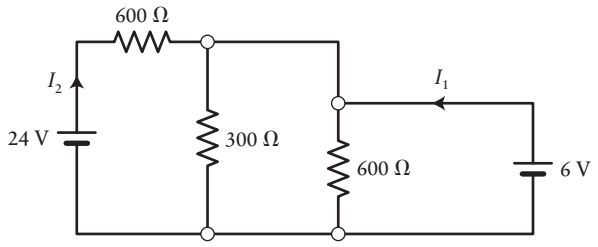
65. For the circuit in Figure P31.65, take the electric potential to be zero at the negative terminal of the battery. Calculate (*a*) the equivalent resistance of the circuit, (*b*) the electric potential at position a, and (*c*) the magnitude and direction of the current through each resistor. ••

Figure P31.65



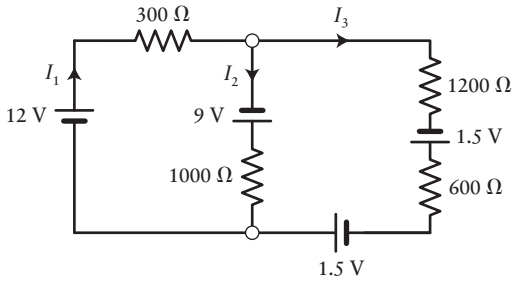
66. In Figure P31.66, determine the magnitude of current I_1. ●●

Figure P31.66



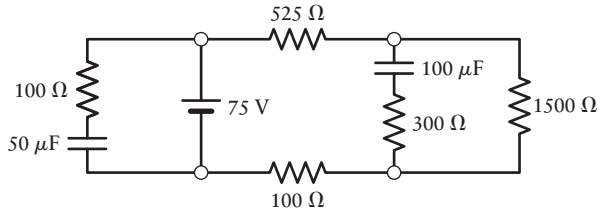
67. In Figure P31.67, calculate the magnitudes of currents I_1, I_2, and I_3. ●●

Figure P31.67



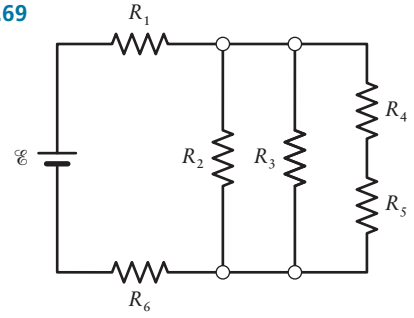
68. In Figure P31.68, the circuit has been completed for several minutes. Calculate (*a*) the current through each resistor and (*b*) the magnitude of charge on each capacitor plate. ●●

Figure P31.68



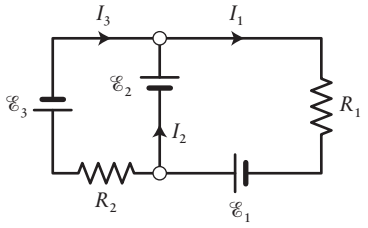
69. In Figure P31.69, calculate (*a*) the equivalent resistance of the circuit and (*b*) the magnitude of the current through each resistor. Use these values: $R_1 = 1.2\ \Omega$, $R_2 = 2.0\ \Omega$, $R_3 = 2.0\ \Omega$, $R_4 = 1.0\ \Omega$, $R_5 = 3.0\ \Omega$, $R_6 = 1.0\ \Omega$, and $\mathscr{E} = 12$ V. ●●

Figure P31.69



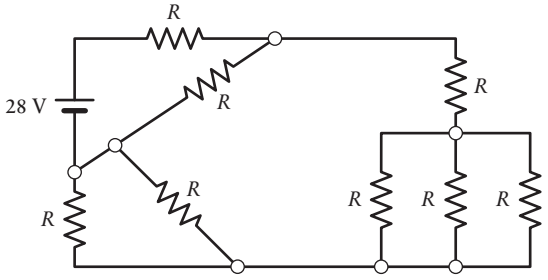
70. In Figure P31.70, determine the magnitudes of the currents I_1, I_2, and I_3 and whether the direction shown for each current is correct or should be reversed. Assume that $R_1 = 3.0\ \Omega$, $R_2 = 6.0\ \Omega$, $\mathscr{E}_1 = 12.0$ V, $\mathscr{E}_2 = 9.0$ V, and $\mathscr{E}_3 = 6$ V. ●●

Figure P31.70



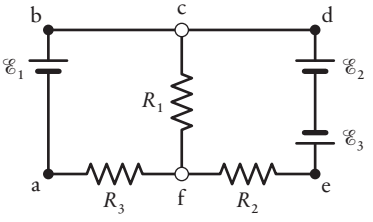
71. The eight resistors in Figure P31.71 are identical to one another, all having resistance $R = 200\ \Omega$. What is the magnitude of the current drawn from the battery? ●●

Figure P31.71



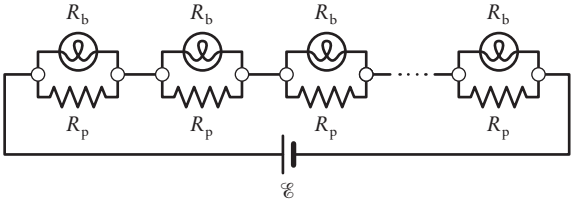
72. In Figure P31.72, $\mathscr{E}_1 = 5.0$ V, $\mathscr{E}_2 = 5.0$ V, $\mathscr{E}_3 = 1.5$ V, $R_1 = 50\ \Omega$, $R_2 = 50\ \Omega$, and $R_3 = 50\ \Omega$. What are (*a*) the current in each branch of the circuit and (*b*) the potential differences V_{ab}, V_{bc}, V_{cd}, V_{de}, V_{ef}, and V_{fa}? ●●

Figure P31.72



73. A string of winter holiday lights consists of N bulbs, each having resistance R_b (Figure P31.73). Wired in parallel with each bulb is a resistor of resistance R_p. What is the resistance of the string? If one of the bulbs burns out, what happens to the other $N - 1$ bulbs? ●●

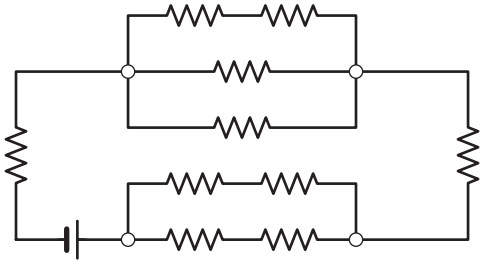
Figure P31.73



74. An ammeter that has internal resistance $R_{am} = 0.504\ \Omega$ is designed to measure a maximum current of $I_{max} = 100$ mA. You want to use this ammeter to measure the current in a circuit that consists of a 3.00-V battery and a resistor that has a resistance of 4.00 Ω. (*a*) What problem do you encounter? (*b*) How can you overcome this problem by inserting a second resistor into the circuit and rescaling the meter readout (that is, by multiplying the readout value by some constant)? ••

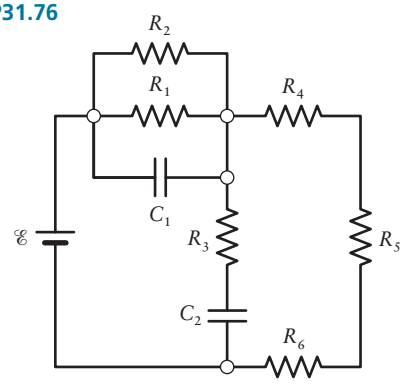
75. If each resistor in Figure P31.75 has resistance $R = 5.0\ \Omega$, what is the equivalent resistance of the combination? ••

Figure P31.75



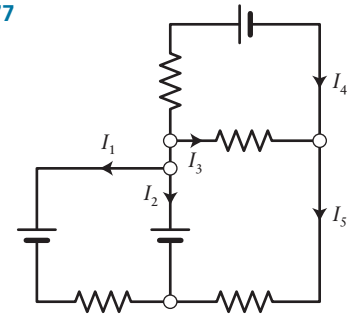
76. In Figure P31.76, $R_1 = 2.0\ \Omega$, $R_2 = 1.5\ \Omega$, $R_3 = 2.0\ \Omega$, $R_4 = 1.0\ \Omega$, $R_5 = 2.0\ \Omega$, $R_6 = 1.0\ \Omega$, $C_1 = 20\ \mu$F, $C_2 = 40\ \mu$F, and $\mathscr{E} = 12.0$ V. Assume that the circuit has been connected for several minutes. (*a*) Calculate the current through the battery and through each resistor. (*b*) Compute the magnitude of charge on each plate of both capacitors. •••

Figure P31.76



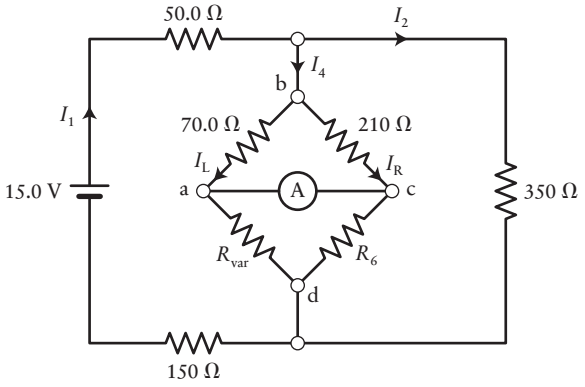
77. In Figure P31.77, each of the three batteries supplies an emf of 6.0 V and each of the four resistors has a resistance of 3.0 Ω. Calculate the magnitudes of the five currents I_1–I_5. •••

Figure P31.77



78. If the ammeter in the Wheatstone bridge of Figure P31.78 measures zero current when the resistance R_{var} of the variable resistor is set to 185 Ω, what is the current I_L through the left side of the bridge? •••

Figure P31.78



31.8 Power in electric circuits

79. If a light bulb has a resistance of 5.5 Ω and is dissipating energy at a rate of 9.0 W, what are (*a*) the current through the bulb and (*b*) the potential difference across its terminals? •

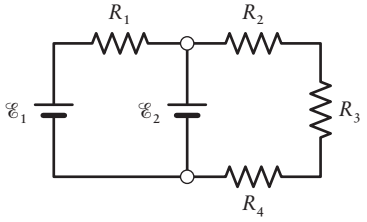
80. Three resistances R, $2R$, and $3R$ are connected in parallel in an electric circuit. Find the ratio of energy dissipation in R, $2R$, and $3R$. •

81. Two light bulbs 1 and 2 are connected in parallel to an 8.00-V battery. (*a*) If the bulb resistances are $R_1 = 4.0\ \Omega$ and $R_2 = 6.0\ \Omega$, what is the rate at which each bulb consumes energy? (*b*) At what rate does the circuit consume energy? (*c*) Calculate the same three power values for the bulbs connected in series rather than in parallel. •

82. A car battery is labeled "12 V 40 Ah." You forget to switch off the light in your glove compartment, which draws 0.80 A. How long until the battery is drained? •

83. (*a*) Determine the current through and the potential difference across each resistor in Figure P31.83. Assume that $R_1 = R_2 = R_3 = R_4 = 3.0\ \Omega$, $\mathscr{E}_1 = 9.0$ V, and $\mathscr{E}_2 = 4.5$ V. (*b*) At what rate is energy dissipated in each resistor? (*c*) What is the power of each battery, and is energy used by or supplied by each battery? ••

Figure P31.83

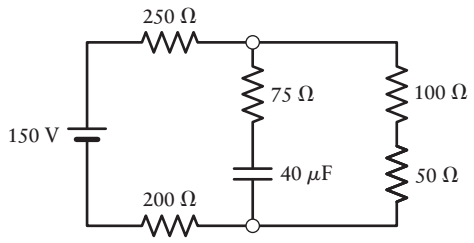


84. The current through a 60-W light bulb is 2 A when connected to a nonideal battery with emf $\mathscr{E} = 120$ V. Determine the bulb's resistance and internal resistance of the battery. ••

85. Determine the current through each resistor and the magnitude of the charge on either capacitor plate after

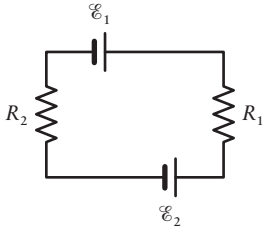
the circuit in Figure P31.85 has been connected for a few minutes. ••

Figure P31.85



86. If the rate at which energy is dissipated by resistor 1 in Figure P31.86 is 2.5 W, and $R_1 = 10\ \Omega$, $\mathscr{E}_1 = 12$ V, and $\mathscr{E}_2 = 6$ V, (*a*) what is the value of R_2? (*b*) At what rate is energy dissipated in resistor 2? (*c*) Which battery supplies energy to the circuit, and at what rate does it do so? (*d*) At what rate is energy used by the other battery? ••

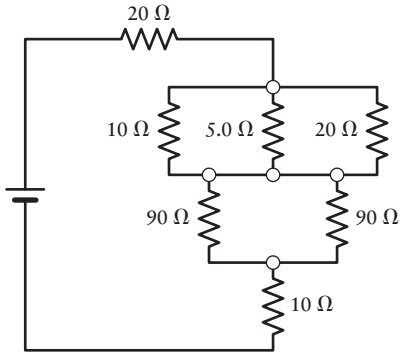
Figure P31.86



87. We have two types of wire made up of copper and aluminum. (*a*) If these wires are of same length, radius, and carry same current, which wire has higher rate of energy loss? (*b*) For same rate of energy loss, length, and radius for both types of wires, calculate the current in aluminum wire, if 15 A of current flows in the copper wire. ••

88. In Figure P31.88, in which resistor is energy dissipated (*a*) at the greatest rate and (*b*) at the smallest rate? •••

Figure P31.88



89. At what rate is energy either delivered by or delivered to each battery in Figure P31.67? At what rate is energy dissipated in each resistor? Does the power summed over all the elements make sense? •••

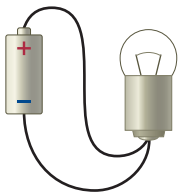
90. A physics student who needs a magnetic field for a project makes a solenoid coil from "magnet wire," which is copper wire coated with a very thin enamel insulation. (The insulation is so thin that you can ignore its thickness.) He chose 28-gauge wire, which is 0.321 mm in diameter, and winds a solenoid that is 50 mm in diameter and 0.20 m long. He then runs a 3.0-A current through the solenoid but is disappointed at how small the magnetic field magnitude is and horrified at how hot the magnet gets after a few minutes. Calculate the magnetic field magnitude and the rate at which energy is lost as thermal energy as the coil heats up. What would you suggest the student do to get a greater magnetic field and less thermal energy loss? •••

Additional Problems

91. Consider a strip of material that has a rectangular cross-section with height 310 μm and width 4.2 mm and which carries a uniform current of 5.5 mA. (*a*) What is the current density in the strip? (*b*) If the number density of the charge carrier (conduction electron) is $3.5 \times 10^{23}\ \text{m}^{-3}$, then calculate the drift speed in the strip. •

92. Does the light bulb in Figure P31.92 light up? Why or why not? •

Figure P31.92

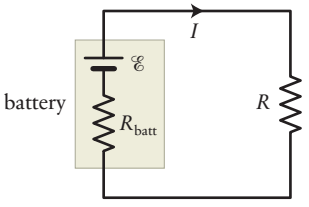


93. Calculate the mean free time τ between the collisions for the conduction electron in copper. [$n = 8.49 \times 10^{28}\ \text{m}^{-3}$, $\rho_{Cu} = 1.68 \times 10^{-8}\ \Omega \cdot \text{m}$] •

94. Given a wire of length 5.5 m and radius 1.2 mm, find the type of material it is made up of if it carries a current of 1.31 A and dissipates electrical energy at the rate of 55.0 mW. •

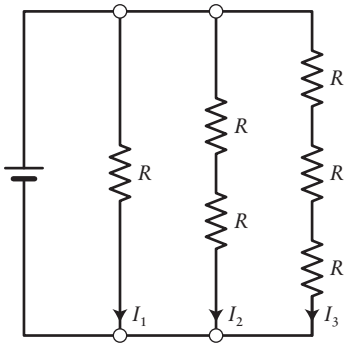
95. The battery in Figure P31.95 has internal resistance $R_{\text{batt}} = 4.0\ \Omega$ and maintains an emf $\mathscr{E} = 12.0$ V. What is the resistance R of the resistor connected in series with the battery if the current through the circuit is 0.5 A? •

Figure P31.95



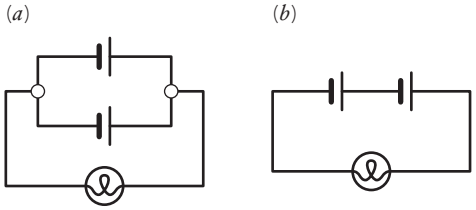
96. Calculate the values of I_1 and I_2 in the circuit shown in Figure P31.96 if all the resistors have resistance 240 Ω and $I_3 = 2.0$ A. ●●

Figure P31.96



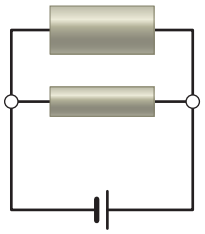
97. If each battery in Figure P31.97 has an emf of 9.0 V, what is the potential difference across the light bulb in each circuit? ●●

Figure P31.97



98. The two resistors in Figure P31.98 are made of the same material and are of equal length. The only difference between them is that the radius of the top one is greater than the radius of the bottom one. (*a*) Through which resistor is the current greater? (*b*) Across which resistor is the potential difference greater? (*c*) How does the resistance of the circuit compare with the individual resistances of the two resistors? ●●

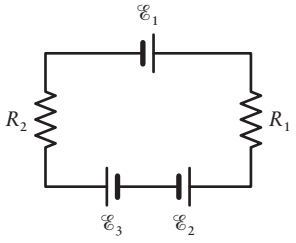
Figure P31.98



99. If you want to add a third resistor to the circuit in Figure P31.98 to reduce the circuit resistance as much as possible, should you connect it in series or in parallel? Should you add a resistor with a large cross-sectional area, like the top one in Figure P31.98, or one with a small cross-sectional area, like the bottom one? ●●

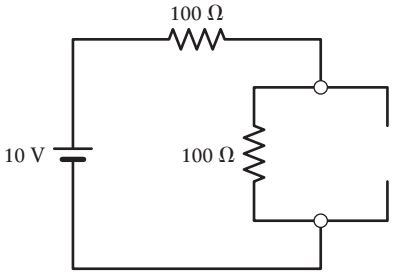
100. In Figure P31.100, $\mathscr{E}_1 = 3.0$ V and $\mathscr{E}_2 = 5.0$ V. (*a*) What value of $\mathscr{E}_3$ causes the potential difference across resistor 1 to be zero? (*b*) In that situation, what is the current through resistor 2? ●●

Figure P31.100



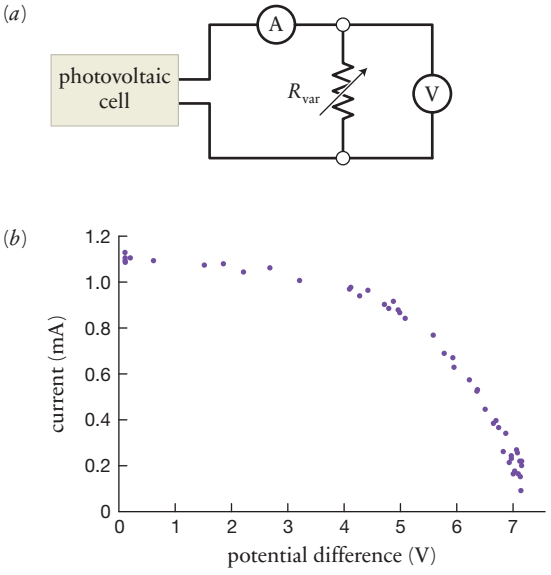
101. Your boss has given you the incomplete circuit shown in Figure P31.101 and charged you with determining the greatest and smallest currents that can be drawn from the battery using this circuit. The catch is that you must close the gap with a resistor from your laboratory. You rummage around and find a huge cabinet full of resistors ranging from very small to very large resistance. ●●●

Figure P31.101



102. You illuminate a photovoltaic cell with a halogen light bulb. The cell is connected to an ammeter, a voltmeter, and a variable resistor (Figure P31.102*a*), and the current through and potential difference across the resistor are measured for different settings of the resistor. Figure P31.102*b* shows the data from these measurements,

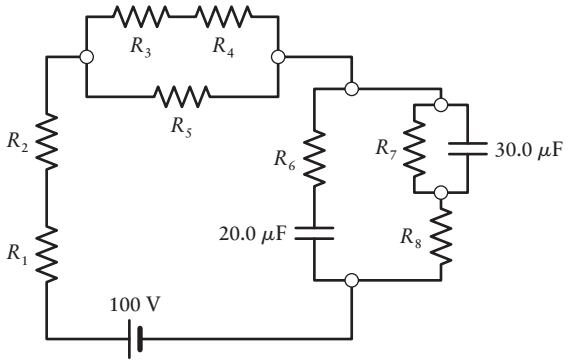
Figure P31.102



where each data point represents the potential difference and current at a particular value of R_{var}. What values of R_{var} do the data points at low and high potential differences correspond to? When is the power of the photovoltaic cell at its maximum value? ●●●

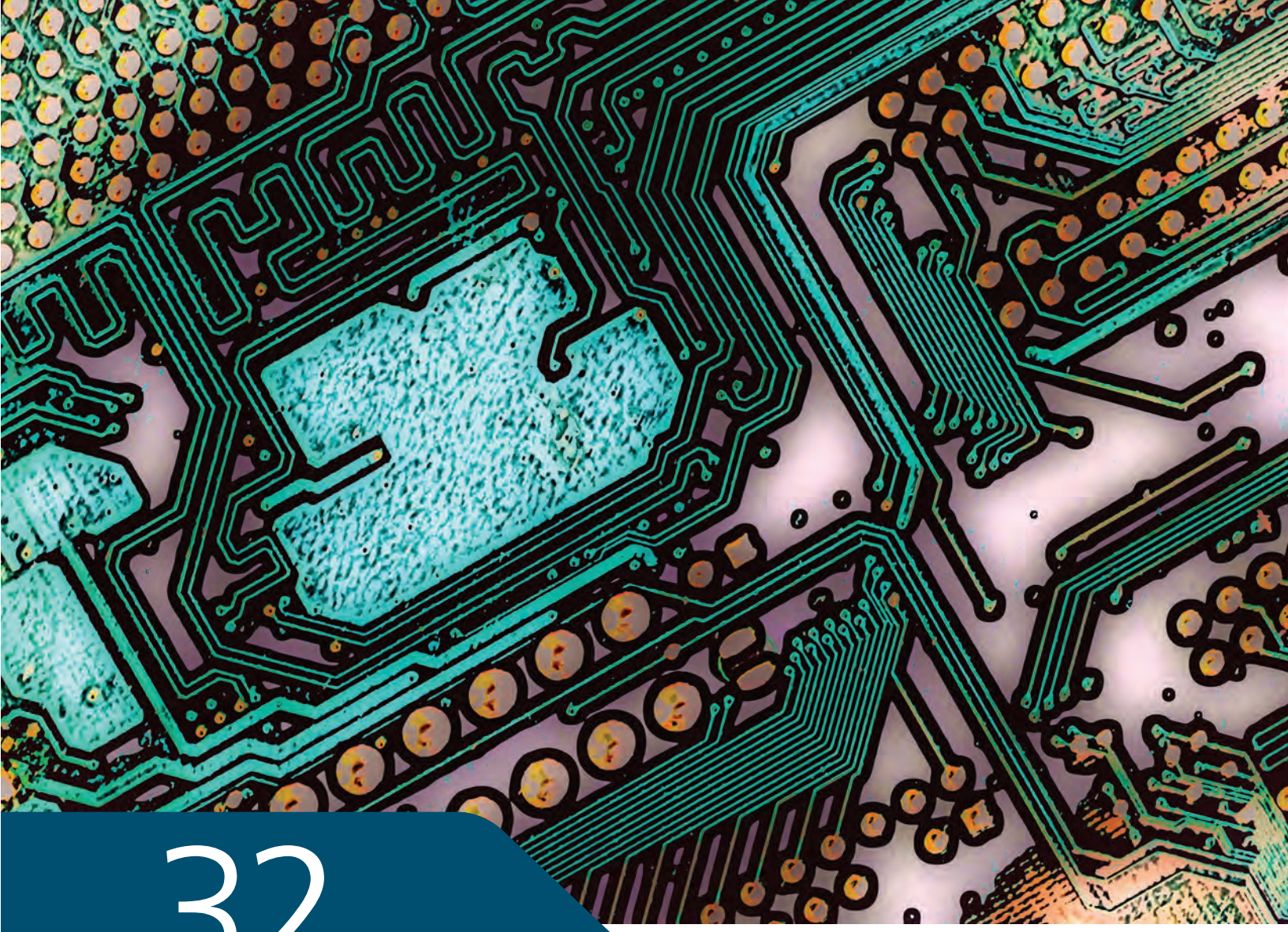
103. The circuit shown in Figure P31.103 has been connected for a few minutes. Determine the current through each resistor and the battery, and determine the magnitude of charge on either plate of each capacitor. Use resistor values $R_1 = R_2 = 5.00\ \Omega$, $R_3 = 4.00\ \Omega$, $R_4 = 6.00\ \Omega$, $R_5 = 10.0\ \Omega$, $R_6 = 0.500\ \Omega$, $R_7 = 1.00\ \Omega$, and $R_8 = 0.500\ \Omega$. ●●●

Figure P31.103



104. Three students rent a third-floor apartment. After signing the lease, they realize there's no electricity on the third floor! They run a 30.5-m, 18-gauge extension cord up from the second floor and plug everything into some power strips plugged into this cord. An electrician visiting the apartment says, "You can't do that; you'll burn the house down! Get a heavy-duty cord made with 12-gauge wire." You don't see how that could make a significant difference, so you begin to look up wire sizes in a reference table. ●●● CR

105. You need a liquid electrical conductor for a project you are working on, and as one possibility you try seawater. You know the number density of the charge carriers and the average time interval between collisions. Your preliminary calculations using Eq. 31.9 and the mass and charge of an electron give you a theoretical conductivity, σ_{theory}. When testing the seawater, though, you obtain a conductivity less than two thousand times smaller than σ_{theory} ($\sigma_{test} < \sigma_{theory}/2000$). You struggle to understand this discrepancy. ●●● CR



32

Electronics

32.1 Alternating currents
32.2 AC circuits
32.3 Semiconductors
32.4 Diodes, transistors, and logic gates

32.5 Reactance
32.6 *RC* and *RLC* series circuits
32.7 Resonance
32.8 Power in AC circuits

CONCEPTS

QUANTITATIVE TOOLS

In the preceding chapter, we discussed electric circuits in which the current is steady. As noted in that chapter, the steady flow of charge carriers in one direction only is called *direct current*. Batteries and other devices that produce static electrical charge, such as van de Graaff generators, are sources of direct current. Although direct current has many uses, it has several limitations as well. For example, in order to produce substantial currents, direct-current sources must be quite large and are therefore cumbersome. More important, steady currents do not generate any electromagnetic waves, which can be used to transmit information and energy through space, as we saw in Chapter 30.

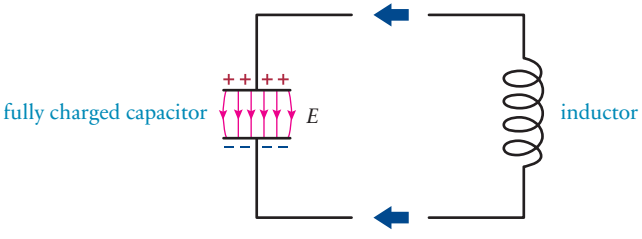
Because of these and many other factors, most electric and electronic circuits operate with **alternating currents** (abbreviated AC)—currents that periodically change direction. The current provided by household outlets in the United States, for instance, alternates in direction, completing 60 cycles per second (that is, with a frequency of 60 Hz), and the currents in computer circuits change direction billions of times per second. It is no understatement to say that contemporary society *depends* on alternating currents.

In this chapter we discuss the basics of both household currents and the electronics that lie at the heart of computers.

32.1 Alternating currents

We have already encountered one example of an electrical device that produces a changing current: a capacitor that is either charging or discharging. Let's consider what happens when we connect an inductor to a charged capacitor (**Figure 32.1**). A circuit that consists of an inductor and a capacitor is called an *LC* circuit. As soon as the two circuit elements are connected, positive charge carriers begin to flow clockwise through the circuit. The magnitude of the current increases from its initial value of zero (**Figure 32.2a**) to a nonzero value (Figures 32.2*b*–*d*). The capacitor discharges through the inductor, and the current causes a magnetic field in the inductor. As the current in the inductor increases, the magnetic field also increases, causing an induced emf (see Section 29.7) that opposes this increase and prevents the current from increasing rapidly. Consequently, the capacitor discharges more slowly than it would if we had connected it to a wire.

Figure 32.1 What happens when we connect an inductor (represented as a coil in circuit diagrams) to a charged capacitor?



32.1 (*a*) Just before the inductor is connected to the charged capacitor, what type of energy is contained in the system comprising the two elements? (*b*) Once the two elements are connected to each other, what happens to that energy? (*c*) Once the capacitor is completely discharged, in what form is the energy in the circuit?

As you saw in Checkpoint 32.1, when the capacitor is completely discharged, all of the energy in an *LC* circuit is contained in the magnetic field and this field reaches its maximum magnitude (Figure 32.2*c*). Because the magnetic energy is proportional to the square of the current in the inductor (Eq. 29.25), the current, too, reaches its maximum value at this instant. Once the magnetic field and the current reach their maximum values, the current begins to charge the capacitor in the opposite direction (Figure 32.2*d*), and the charge on the capacitor increases as the magnetic field in the inductor decreases. When the magnetic field in the inductor is zero, the current is also zero and the capacitor has again maximum charge but with the opposite polarity (Figure 32.2*e*). The process then repeats itself with the current in the opposite direction (Figures 32.2*f*–*h*) until the capacitor is restored to its starting configuration. Then the cycle begins again.

Figure 32.3 shows the time dependence of the electric potential energy U^E stored in the capacitor and the magnetic potential energy U^B stored in the inductor. In the absence of dissipation, the energy in the circuit, $U^E + U^B$, must stay constant. Therefore, when the capacitor is not charged and U^E drops to zero, U^B must reach its maximum value, U_{max}.

There is always some dissipation in a circuit. Resistance in the connecting wires gradually converts electrical energy to thermal energy. Consequently, the oscillations decay in the same manner as the damped mechanical oscillations we considered in Section 15.8. Resistance therefore plays the same role in oscillating circuits as damping does in mechanical oscillators.

Throughout this chapter we work with time-dependent potential differences and currents. To make the notation as concise as possible, we represent time-dependent quantities with lowercase letters. In other words, v_C is short for $V_C(t)$ and i is short for $I(t)$. We also need a symbol for the maximum value of an oscillating quantity—its *amplitude* (see Section 15.1). For this we use a capital letter without the time-dependent marker (t); thus V_C is the maximum value of the potential difference across a capacitor, and I is the maximum value of the current in a circuit.

Unlike their counterparts in DC circuits, the potential difference across the capacitor, v_C, and the current in the *LC* circuit, i, change sign periodically. So, when analyzing AC circuits, we must carefully define what we mean by the sign of these quantities. To analyze the *LC* circuit in Figure 32.2, for example, we choose a reference

CONCEPTS

Figure 32.2 A series of "snapshots" showing what happens when we connect an inductor to a charged capacitor.

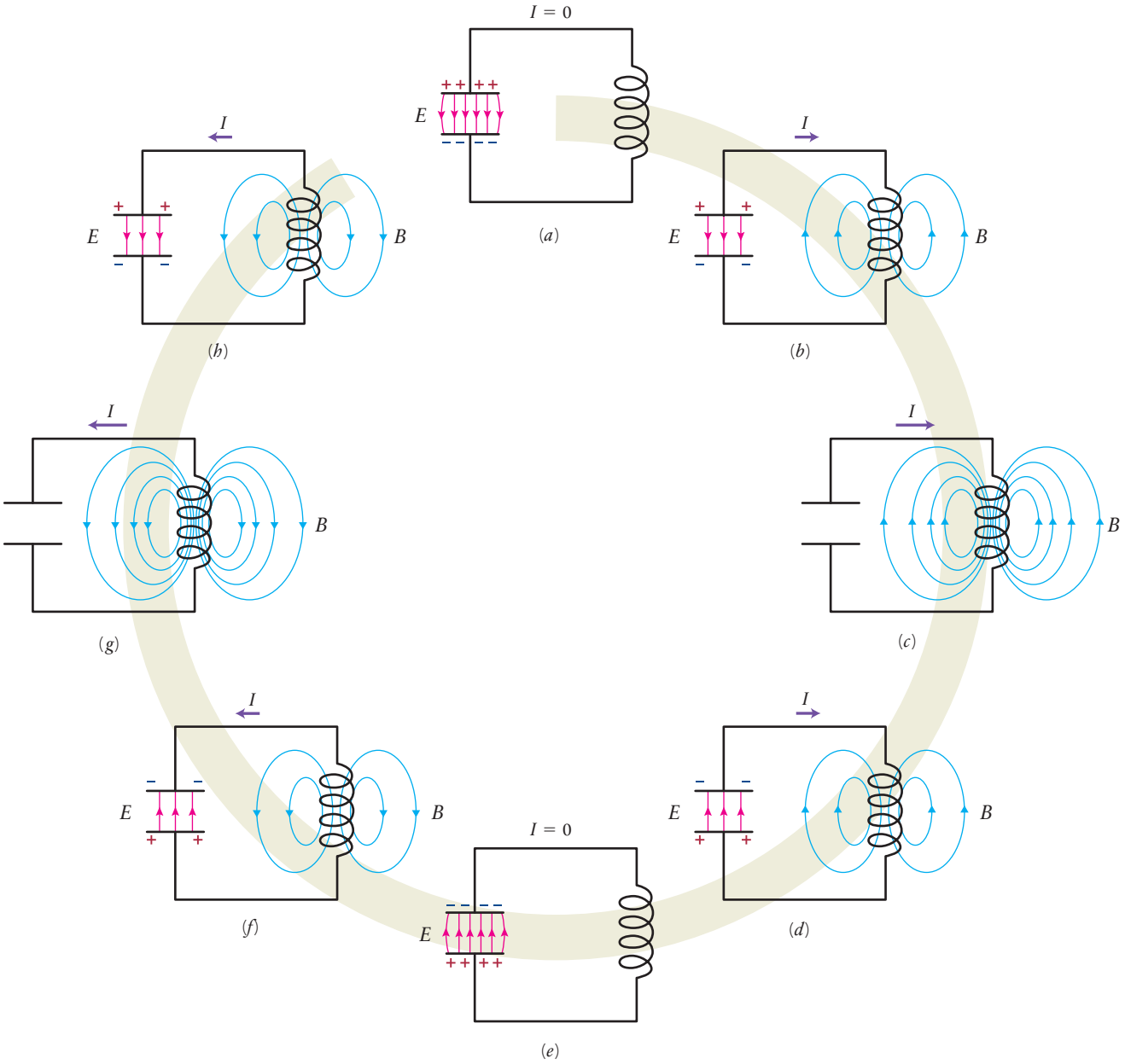
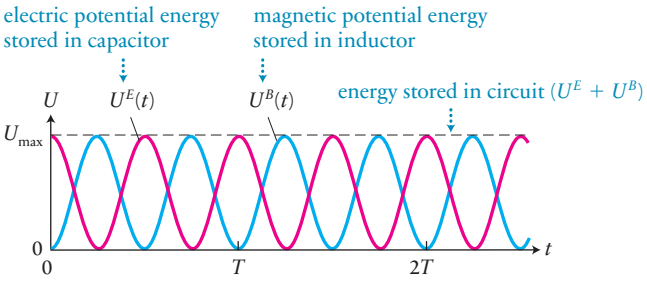


Figure 32.3 Time dependence of the electric potential energy U^E stored in the capacitor and the magnetic potential energy U^B stored in the inductor. In the absence of dissipation, the energy in the circuit, $U^E + U^B$, is a constant U_{max}.

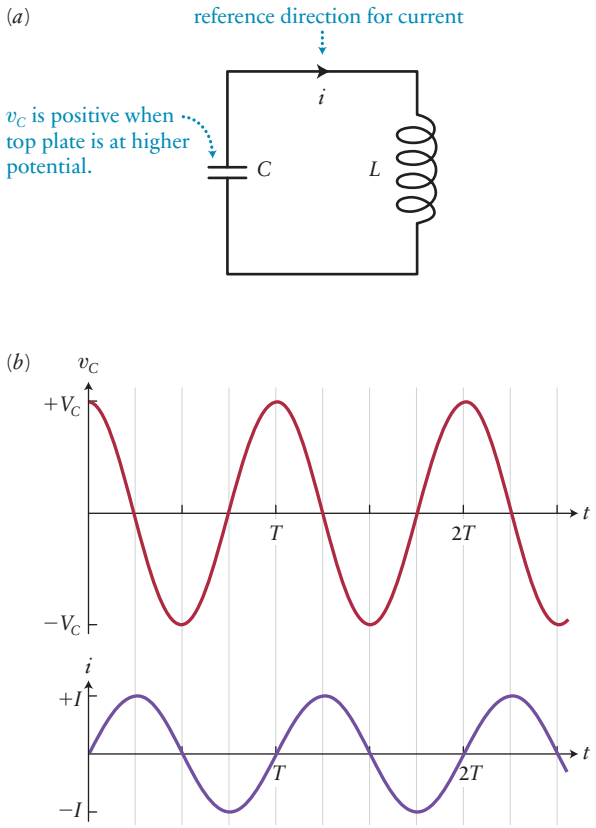


direction for the current i and let the potential difference v_C be positive when the top capacitor plate is at a higher potential than the bottom plate (**Figure 32.4*a*** on the next page). Note that both of these choices are arbitrary.

Figure 32.4*b* shows graphs for v_C and i with these choices. The potential difference across the capacitor v_C is initially positive, representing the situation at Figure 32.2*a*. During the first quarter cycle (Figure 32.2*b*), the capacitor is discharging and positive charge carriers travel away from the top plate of the capacitor in the chosen reference direction, and so i is positive. In the part of the cycle represented by Figure 32.2*f*, where the capacitor is again discharging, v_C is negative (because the top plate is negatively

CONCEPTS

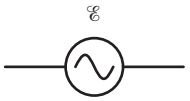
Figure 32.4 For the *LC* circuit shown in Figure 32.2, graphs of the time-dependent potential difference across the capacitor (defined to be positive when the top plate is at the higher potential) and the current in the circuit (defined to be positive when positive charge carriers travel away from the top plate of the capacitor). One cycle is completed in a time interval T (the *period*).



charged) and i is negative (because the direction of current is opposite the chosen reference direction), as shown in the time interval $\frac{1}{2}T < t < \frac{3}{4}T$ in Figure 32.4*b*. (See if you can work out the signs during the time intervals when the capacitor is charging.) Both v_C and i vary sinusoidally in time, with v_C at its maximum when i is zero, and vice versa.

Because of dissipation, the *LC* circuit in Figure 32.1 is not a practical source of alternating current; instead, generators are widely used to produce sinusoidally alternating emfs in a circuit (see Example 29.6). The symbol for a source that generates a sinusoidally alternating potential difference and/or current is shown in **Figure 32.5**; such a source is called an **AC source**. The time-dependent emf an AC source produces across its terminals is designated $\mathscr{E}$, and its amplitude is designated $\mathscr{E}_{max}$.

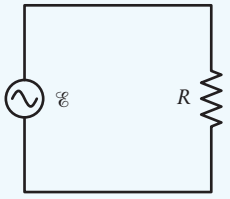
Figure 32.5 Symbol that represents an AC source in an electric circuit. The AC source produces a sinusoidally varying emf $\mathscr{E}$ across its terminals.



EXERCISE 32.1 AC source and resistor

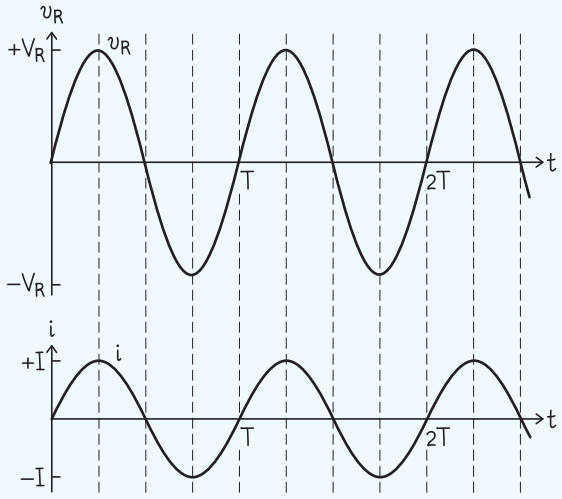
Figure 32.6 shows a circuit consisting of an AC source and a resistor. The emf produced by the generator varies sinusoidally in time. Sketch the potential difference across the resistor as a function of time and the current in it as a function of time.

Figure 32.6 Exercise 32.1.



SOLUTION Ohm's law, the junction rule, and the loop rule (see Chapter 31) apply to alternating-current circuits just as they do to direct-current circuits. All I need to remember here is that the potential differences and currents are time dependent. Applying the loop rule to this circuit requires the time-dependent potential difference across the resistor v_R to equal the emf $\mathscr{E}$ of the AC source at every instant, so that the sum of the potential differences around the circuit is always zero. Consequently, v_R oscillates just as $\mathscr{E}$ oscillates, as shown in **Figure 32.7**; V_R is the maximum value of the potential difference across the resistor.

Figure 32.7



Ohm's law requires the time-dependent current i in the resistor to be proportional to v_R, which means that i also oscillates, with the current at its maximum when v_R is maximum. ✔

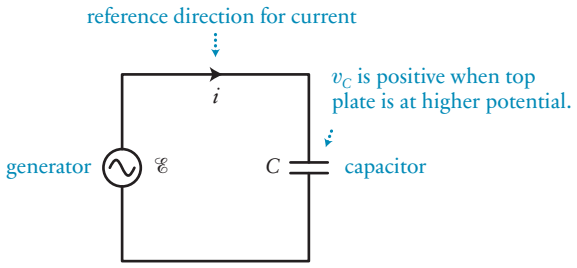
32.2 (*a*) Is energy dissipated in the resistor in the circuit of Figure 32.6? (*b*) If so, why doesn't the amplitude of the oscillations of v_R and i (shown in Figure 32.7) decrease with time?

32.2 AC circuits

The circuit discussed in Exercise 32.1 is an alternating current, or AC circuit. Such circuits exhibit more complex behavior when they contain elements that do not obey Ohm's law, so that the current is not proportional to the emf of the source. For example, let's consider the current in the circuit shown in **Figure 32.8**.

CONCEPTS

Figure 32.8 AC circuit with a capacitor connected to an AC source.



To analyze the circuit we choose a reference direction for the current i and let the potential difference v_C again be positive when the top capacitor plate is at a higher potential than the bottom plate (Figure 32.8). Because the capacitor is connected directly to the AC source, the time-dependent potential difference across the capacitor v_C equals the emf of the AC source at any instant. What is the current in the circuit? Let's begin considering what happens when the capacitor is uncharged. As v_C increases, the charge on the top plate of the capacitor increases. This means that positive charge carriers are moving toward the top plate, in the same direction as the chosen reference direction for the current, and so the current is positive (**Figure 32.9*a***). When v_C reaches its maximum, the capacitor reaches its maximum charge and the current is instantaneously zero. As v_C decreases, the charge on the top plate of the capacitor decreases. Positive charge carriers now move away from the top plate and the current is negative (Figure 32.9*b*). At some instant the top plate becomes negatively charged (Figure 32.9*c*); v_C continues to decrease until it reaches its minimum value and the current is instantaneously zero. At that instant the capacitor again reaches its maximum charge but with the opposite polarity. As v_C begins to increase again, positive charge carriers flow toward the top plate and the current is positive again (Figure 32.9*d*). When both plates are uncharged again, the cycle is complete.

Figure 32.9 The charging and discharging of the capacitor in the circuit of Figure 32.8.

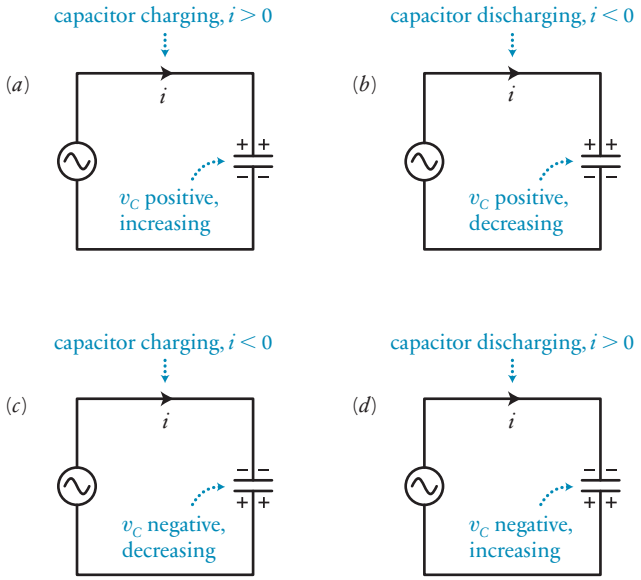


Figure 32.10 Time-dependent current in the circuit and potential difference across the capacitor for the circuit of Figure 32.9.

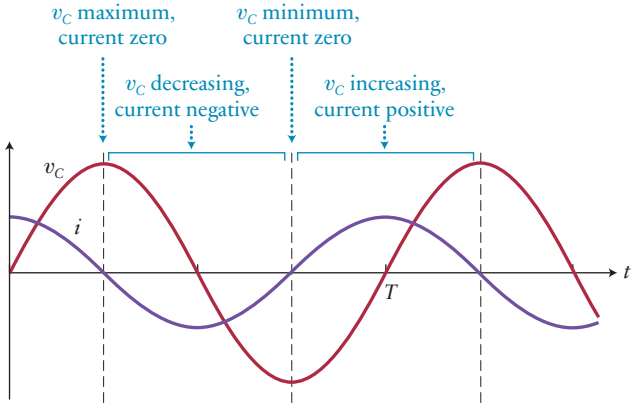


Figure 32.10 shows the time dependence of i and v_C in Figure 32.9. Note that i and v_C are not simply proportional to one another. Instead, the current maximum occurs one-quarter cycle before the potential difference maximum. For this reason, the current is said to *lead* the potential difference:

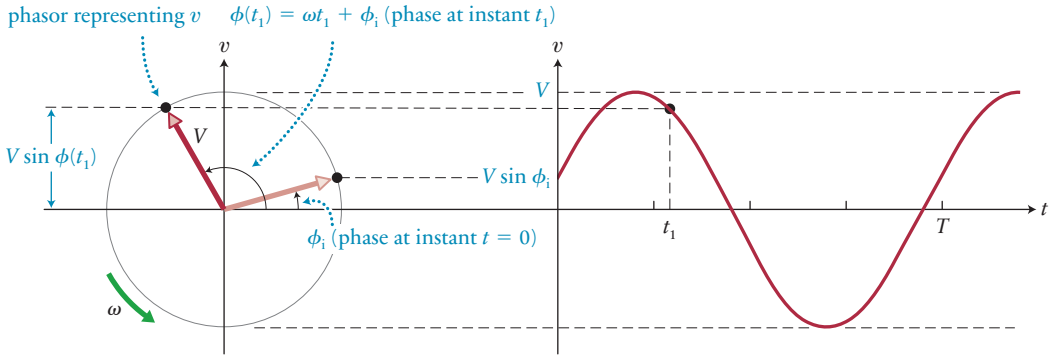
> **In an AC circuit that contains a capacitor, the current in the capacitor leads the potential difference by 90° (a quarter of an oscillation cycle).**

To describe the time dependence of a sinusoidally oscillating quantity, we must specify both the angular frequency of oscillation ω and the instant at which the oscillating quantity equals zero. As discussed in Chapter 15, a sinusoidally time-dependent quantity (such as the circuit potential difference we are looking at here) can be written in the form $v = V\sin(\omega t + \phi_i)$. The argument of the sine, $\omega t + \phi_i$, is the *phase*. At $t = 0$ the phase is equal to the *initial phase* ϕ_i (Chapter 15). When the phase of an oscillating quantity is zero, $\omega t + \phi_i = 0$, the quantity is zero as well because $\sin(0) = 0$.

We can analyze phase differences in AC circuits with lots of algebra, but the underlying physics is much clearer (and the analysis much simpler!) if we use the phasor notation developed in Chapter 15 to describe oscillatory motion. Following the approach of Section 15.5, we can represent an oscillating potential difference v by a phasor rotating in a reference circle (**Figure 32.11** on the next page). Because the length of the phasor equals the amplitude (maximum value) of v, the phasor is labeled V. The phasor rotates counterclockwise at angular frequency ω. The magnitude of v at any instant is given by the vertical component of the phasor; as the phasor rotates, that component oscillates sinusoidally in time, as shown in Figure 32.11. The angle measured counterclockwise from the positive horizontal axis to the phasor is the phase $\omega t + \phi_i$.

CONCEPTS

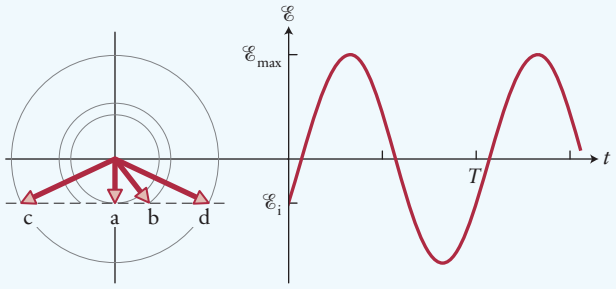
Figure 32.11 Phasor representation of a sinusoidally varying potential difference v. The phasor rotates counterclockwise at the same angular frequency at which v oscillates. The instantaneous value of v equals the length of the vertical component of the phasor.



EXAMPLE 32.2 Phasors

Consider the oscillating emf represented in the graph of **Figure 32.12**. Which of the phasors a–d, each shown at $t = 0$, correspond(s) to this oscillating emf?

Figure 32.12 Example 32.2.



1 GETTING STARTED I begin by observing from the graph that the emf is negative at instant $t = 0$ and increases until it reaches a maximum value $\mathscr{E}_{max}$.

2 DEVISE PLAN To identify the correct phasor or phasors, I can use the following information: (1) the length of the phasor is equal to the amplitude of the oscillation, (2) the value of the emf at any instant corresponds to the vertical component of the phasor, and (3) the phasor rotates counterclockwise around the reference circle.

3 EXECUTE PLAN The amplitudes of phasors a and b are too small and so I can rule these two out. The fact that the emf starts out negative at $t = 0$ and then increases tells me that the phasor representing it must be in the fourth quadrant (below the horizontal axis and to the right of the vertical axis), meaning the correct phasor must be d. ✔

4 EVALUATE RESULT I can verify my answer by tracing out the projection of the phasor on the vertical axis as the phasors rotates counterclockwise. The initial value of the projection, initial phase, and amplitude all agree with the values of these variables represented in the graph.

32.3 Construct a phasor diagram for the time-dependent current and potential difference at $t = 0$ in the AC source-resistor circuit of Figure 32.6.

We can generalize the result of this checkpoint to represent i and v_R from Figure 32.7 at an arbitrary instant t_1. Because i and v_R are in phase for a resistor, the two phasors for i and v_R always have the same phase and so overlap (**Figure 32.13**). Note that the initial phase ϕ_i is zero because i and v_R are zero at $t = 0$ (at that instant both phasors point to the right along the horizontal axis).

The relative lengths of the I and V_R phasors are meaningless because the units of i and v_R are different. However, for circuits with multiple elements (resistors, inductors, or capacitors), the relative lengths of phasors showing the potential differences across different elements are meaningful and will prove very useful in analyzing the circuit.

Phasors are most useful when we need to represent quantities that are not in phase. **Figure 32.14** shows the phasor diagram that corresponds to Figure 32.10 (at the instant represented by Figure 32.9*a*). As the phasor

Figure 32.13 Phasor diagram and graph showing time dependence of v_R and i from Figure 32.7.

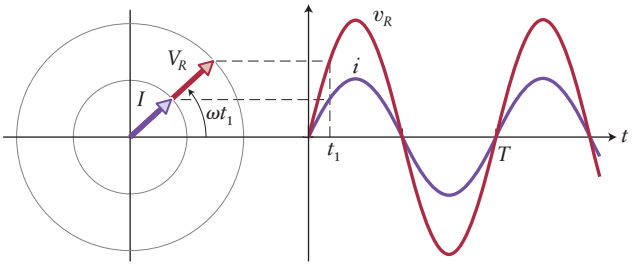
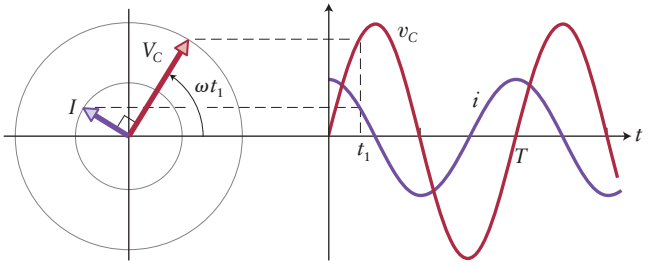


Figure 32.14 Phasor diagram and graph showing time dependence of i and v_C corresponding to Figure 32.10.



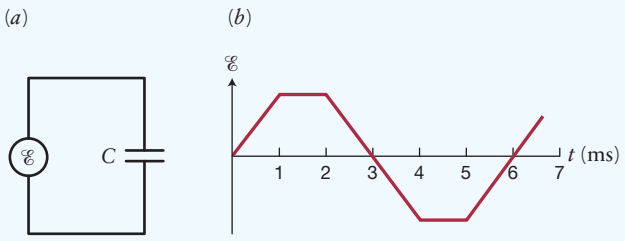
CONCEPTS

diagram shows, the angle between V_C and I is 90°, and so the phase difference between the two phasors is $\pi/2$. Because the phasors rotate counterclockwise, we see that current phasor I is ahead of the potential difference phasor V_C, in agreement with our earlier conclusion that the current in a capacitor leads the potential difference across the capacitor.

EXAMPLE 32.3 Nonsinusoidal AC circuit

When a certain capacitor is connected to a nonsinusoidal source of emf as in **Figure 32.15a**, the emf varies in time as illustrated in Figure 32.15*b*. Sketch a graph showing the current in the circuit as a function of time.

Figure 32.15 Example 32.3.

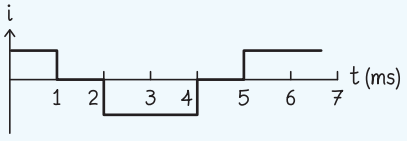


❶ **GETTING STARTED** From Figure 32.15*b* I see that the emf has five distinct parts during the time interval shown. During each part, the emf either is changing at a constant rate or is constant.

❷ **DEVISE PLAN** I know that the current is proportional to the rate at which the charge on the capacitor plates changes over time. I also know that the emf is proportional to the charge on the plates, and so the current is proportional to the derivative of the emf with respect to time.

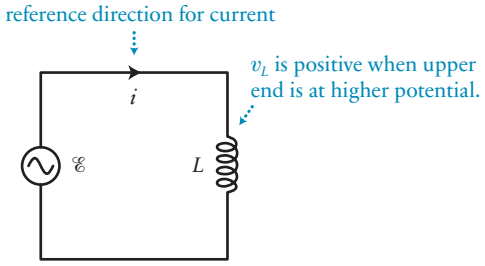
❸ **EXECUTE PLAN** Between $t = 0$ and $t = 1$ ms, the emf increases at a constant rate, so $i = Cd\mathscr{E}/dt$ is constant and positive. Between $t = 1$ ms and $t = 2$ ms, the emf is constant, so $i = Cd\mathscr{E}/dt = 0$. Between $t = 2$ ms and $t = 4$ ms, the emf decreases at a constant rate, so $i = Cd\mathscr{E}/dt$ is constant and negative. Because the rate of decrease between $t = 2$ ms and $t = 4$ ms is the same as the rate of increase between $t = 0$ and $t = 1$ ms, the magnitude of the current between $t = 2$ ms and $t = 4$ ms should be the same as that between $t = 0$ and $t = 1$ ms. The current is zero again during the next millisecond ($t = 4$ ms to $t = 5$ ms) because here the emf is again constant. After $t = 5$ ms, the emf increases again at the same constant rate as between $t = 0$ and $t = 1$ ms, so the current has the same positive value as between $t = 0$ and $t = 1$ ms. The graph representing these current changes is shown in **Figure 32.16**. ✔

Figure 32.16



❹ **EVALUATE RESULT** When the current is positive, the emf is increasing; when the current is negative, the emf is decreasing; and when the current is zero, the emf is constant, as it should be.

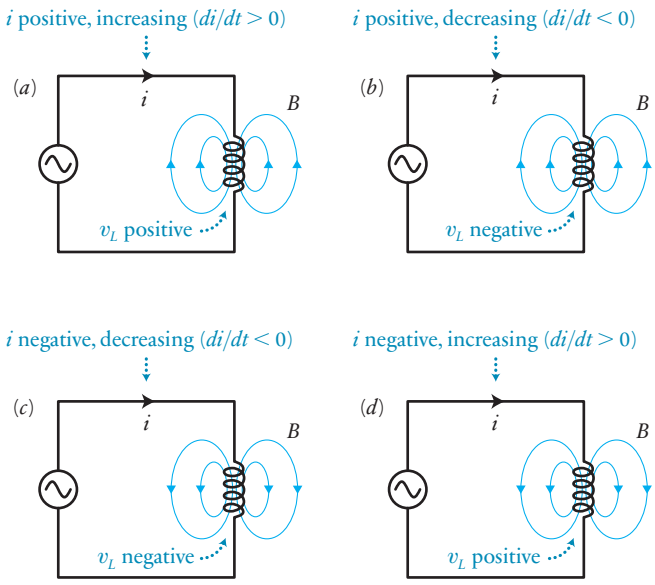
Figure 32.17 AC circuit consisting of an inductor connected across the terminals of an AC source.



Now let's examine the behavior of an inductor connected to an AC generator (**Figure 32.17**). When the current in the circuit is changing, an emf is induced in the coil, in a direction to oppose this change (see Section 29.7). The potential difference between the ends of the inductor, which we'll denote by v_L, is proportional to the rate di/dt at which the current changes (Eq. 29.19). If the current is increasing in the reference direction for current indicated in Figure 32.17, the upper end of the inductor must be at a higher potential than the lower end to oppose the increase in current. If we take v_L to be positive when the upper end of the coil is at a higher potential, v_L must therefore be positive when the current is increasing in the reference direction for the current. This situation is represented in **Figure 32.18a**.

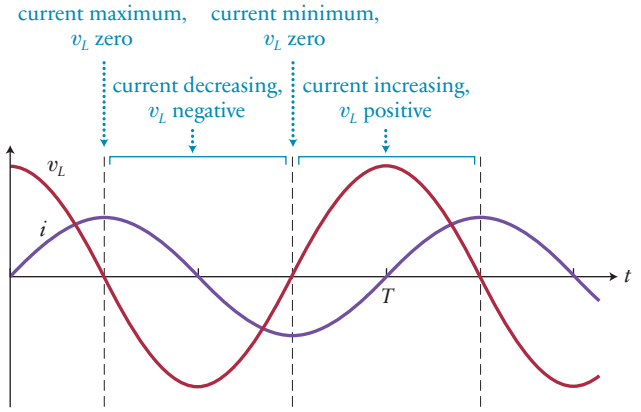
When the current reaches its maximum value in the cycle, v_L is instantaneously zero. After this instant, the current begins to decrease and the lower end of the inductor must be at a higher potential than the upper end to oppose this decrease in current. The potential difference v_L is now negative (Figure 32.18*b*). In the

Figure 32.18 Current and magnetic field oscillations through the inductor of Figure 32.17.



CONCEPTS

Figure 32.19 Graph of time-dependent current in the circuit and potential difference across the inductor for the circuit in Figure 32.17.



second half of the cycle, the current is in the opposite direction. As in the first part of the cycle, v_L has the same sign as di/dt (Figures 32.18*c* and *d*).

Figure 32.19 illustrates the time dependence of i and v_L in Figure 32.18. Note that the current maximum occurs one-quarter cycle after the potential difference maximum. For this reason, the current is said to *lag* the potential difference:

> **In an AC circuit that contains an inductor, the current in the inductor lags the potential difference by 90°.**

Figure 32.20 shows the phasor diagram that corresponds to Figure 32.19 (at the instant represented by Figure 32.18*a*). Just as with the capacitor, the angle between V_L and I is 90° and so the phase difference is $\pi/2$, but in this case the current phasor I is behind the potential difference phasor V_L, in agreement with our earlier conclusion that the current in an inductor lags the potential difference across the inductor.

32.4 What are the initial phases for the phasors in Figures 32.13 and 32.20?

Figure 32.20 Phasor diagram and graph showing time dependence of i and v_L corresponding to Figure 32.19.

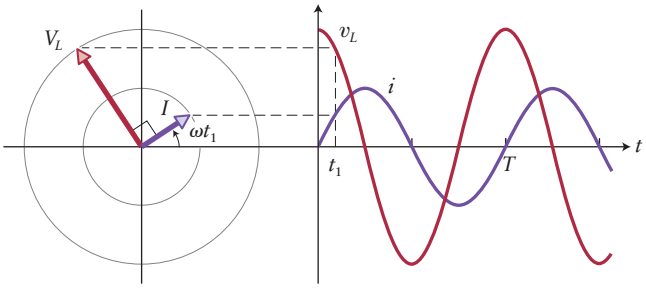
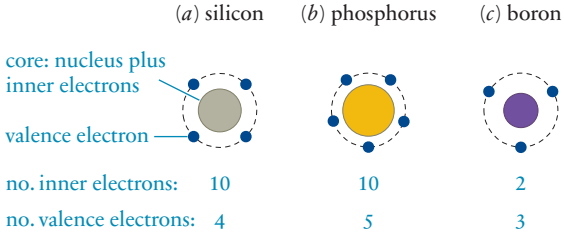


Figure 32.21 Schematic depiction of silicon, phosphorus, and boron atoms, shown as an inner core surrounded by valence electrons.

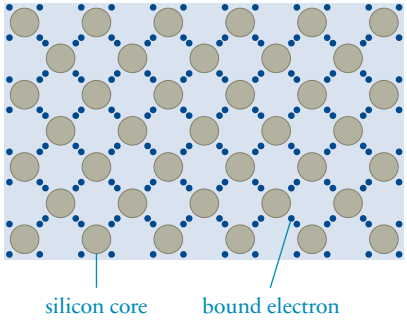


32.3 Semiconductors

Most modern electronic devices are made from a class of materials called **semiconductors**. Semiconductors have a limited supply of charge carriers that can move freely; consequently, their electrical conductivity is intermediate between that of conductors and that of insulators. Semiconductors are widely used in the manufacture of electronic devices such as transistors, diodes, and computer chips because their conductivity can be tailored chemically for particular applications layer by layer, even within a single piece of semiconductor.

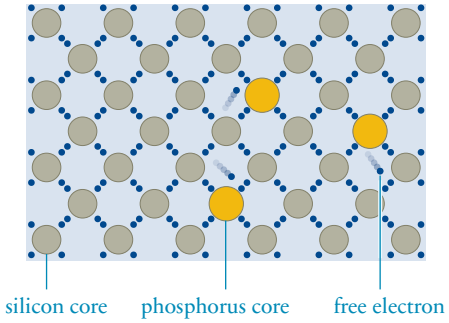
Semiconductors are of two main types: intrinsic and extrinsic. *Intrinsic semiconductors* are chemically pure and have poor conductivity. *Extrinsic* or *doped semiconductors* are not chemically pure, have a conductivity that can be finely tuned, and are widely used in the microelectronics industry. The most widely used semiconductor is silicon, a nonmetallic element that makes up more than one-quarter of Earth's crust. **Figure 32.21*a*** shows a schematic of a silicon atom, which consists of a nucleus surrounded by fourteen electrons. Ten of these electrons are tightly bound to the nucleus—we'll refer to these electrons plus the nucleus as the *core* of the atom. The remaining outermost four electrons are called the atom's *valence electrons*. Each valence electron can form a covalent bond with a valence electron of another silicon atom. These bonds hold many identical silicon atoms together in a crystalline lattice (**Figure 32.22**). The electrons in a covalent bond are not free to move; consequently, pure

Figure 32.22 Schematic of a crystalline lattice of silicon atoms, showing electrons participating in silicon-silicon bonds. (A real silicon crystal exists in three dimensions, and not all of the silicon-silicon bonds lie in a plane; this diagram illustrates only the essential idea that all of the valence electrons participate in covalent bonds.)



CONCEPTS

Figure 32.23 Schematic depiction of a crystalline lattice of silicon atoms doped with phosphorus atoms. The only charge carriers that are free to move in the crystal are the free electrons supplied by the phosphorus dopant atoms.



silicon has a very low electrical conductivity because all of its valence electrons form covalent bonds.

In extrinsic silicon, other types of atoms, such as boron or phosphorus, replace some of the atoms in the silicon lattice, introducing freely moving charge carriers into the lattice. The substituted atoms are called either *impurities* or *dopants*. For example, phosphorus has five valence electrons (Figure 32.21*b*). Because the silicon lattice structure requires only four bonds from each atom, the fifth electron from a phosphorus atom dopant is not involved in a bond and is free to move through the solid (**Figure 32.23**).

If an electric field is applied to the doped semiconductor of Figure 32.23, the free electrons move, creating a current in the semiconductor (**Figure 32.24**). As free electrons leave the semiconductor from one side, other free electrons enter it on the opposite side. Because the semiconductor must remain electrically neutral, the number of free electrons in the semiconductor at any given instant is always the same and it is equal to the number of phosphorus atoms in the material.

If boron atoms, which have three valence electrons (Figure 32.21*c*), are substituted for some silicon atoms

Figure 32.24 In an applied electric field, the free electrons in a phosphorus-doped silicon semiconductor can move in the direction opposite the field direction. Free electrons leave the semiconductor at the left, travel through the circuit wire, and re-enter at the right.

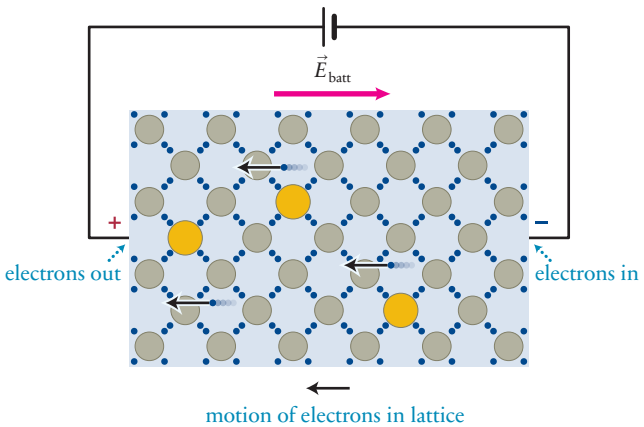
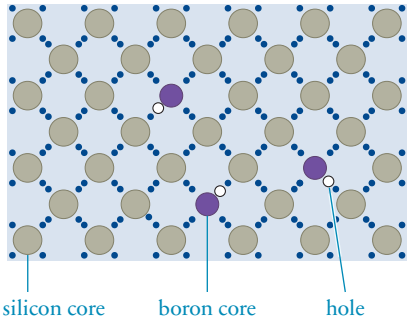


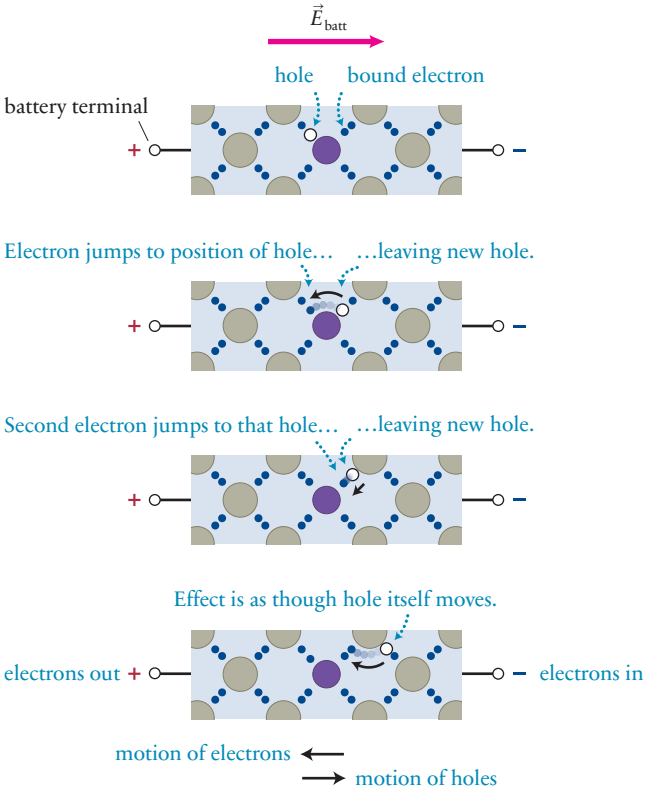
Figure 32.25 Schematic of crystalline lattice of silicon atoms with some boron atoms substituted for silicon, showing both bonding electrons and holes (missing electrons). The only free charge carriers in the crystal are the holes caused by the boron impurities.



in a silicon lattice, the "missing" fourth electron at each boron leaves behind what is called a **hole**—an incomplete bond (**Figure 32.25**). These holes behave like *positive* charge carriers and are free to move through the lattice. The holes therefore increase the ability of the silicon to conduct current, just as do the free electrons in phosphorus-doped silicon.

Keep in mind that the motion of holes involves electrons moving to fill existing holes, leaving new holes in the previous positions of the electrons (**Figure 32.26**).

Figure 32.26 Sequence of four snapshots showing how holes "move" through a crystal by trading places with bonding electrons. In the presence of an electric field, holes move in the direction of the field (opposite to the directions in which the electrons move). To maintain continuity, free electrons from attached metal wires enter at the right, recombining with holes that accumulate there, and leave at the left.



CONCEPTS

The boron cores do not move! In the presence of an electric field, the positively charged holes move in the direction of the field as the negatively charged electrons move in the opposite direction. If the semiconductor is attached to metal wires on either side, as in Figure 32.26, free electrons travel into the semiconductor from the right (eliminating holes that reach the right edge) and travel out of the semiconductor on the left (producing holes on the left edge). Electrons thus flow from right to left, making holes travel in the opposite direction. Unlike the electrons, however, the holes never leave the semiconductor.

Doped semiconductors are classified according to the nature of the dopant. In a *p-type* semiconductor, the dopant has fewer valence electrons than the host atoms, contributing positively charged holes as the free charge carriers (thus the *p* in the name). In an *n-type* semiconductor, the dopant has more valence electrons than the host atoms, contributing negatively charged electrons as the free charge carriers (thus the *n* in the name). Substituting as few as ten dopant atoms per million silicon atoms produces conductivities appropriate for most electronic devices.

32.5 Is a piece of *n*-type silicon positively charged, negatively charged, or neutral?

32.4 Diodes, transistors, and logic gates

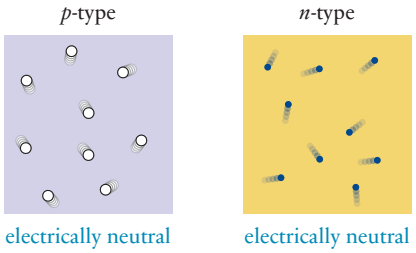
Although tailoring the conductivity of a single piece of semiconductor can be a useful procedure, the most versatile semiconductor devices combine doped layers that have different types of charge carriers. The simplest such device is a **diode**, made by bringing a piece of *p*-type silicon into contact with a piece of *n*-type silicon (**Figure 32.27a**). Near the junction where the two pieces meet, free electrons from the *n*-type silicon wander into the *p*-type material, where they end up filling holes. This *recombination* process turns free electrons into bound electrons (that is, electrons not free to roam around in the material) and eliminates the holes. Likewise, some of the holes in the *p*-type silicon wander into the *n*-type silicon, where they recombine with free electrons.

As recombination events take place, a thin region containing no free charge carriers (neither free electrons nor holes), called the **depletion zone**, develops at the junction. Although there are no *free* charge carriers in this zone, the trapping of electrons on the *p*-side of the junction causes negative charge carriers that are nonmobile to accumulate there. Similarly, positive nonmobile charge carriers accumulate on the *n*-side of the junction. As a result, the depletion zone consists of a negatively charged region and a positively charged region, and an electric field points across the depletion zone from the *n*-side to the *p*-side (Figure 32.27*b*).

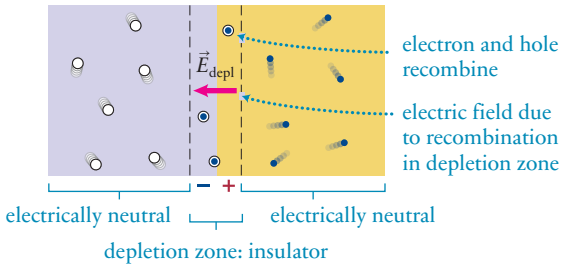
CONCEPTS

Figure 32.27 How a diode transmits current in one direction but blocks it in the other. If the battery is connected as shown in part *d* and produces a sufficiently strong electric field to compensate for the electric field of the depletion zone, there is a steady flow of both electrons and holes. (Remember, though: The holes never leave the semiconductor. Only the electrons enter and leave the semiconductor.)

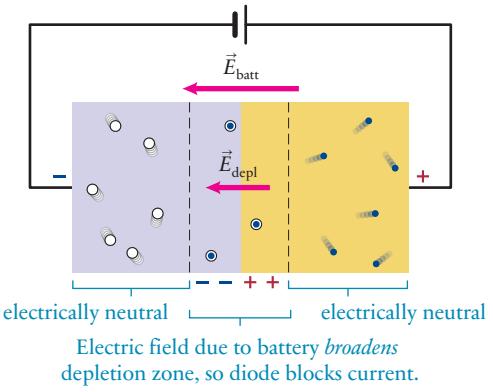
(*a*) Pieces of *p*- and *n*-type doped silicon



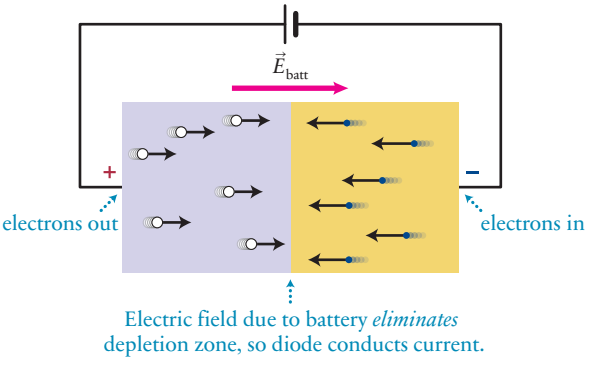
(*b*) When the two are put in contact, a diode is formed



(*c*) Battery connected so as to produce electric field in *same* direction as electric field in depletion zone; diode blocks current



(*d*) Battery connected with the *opposite* polarity; diode conducts current



As this electric field in the depletion zone of the diode increases, it becomes more difficult for free electrons and holes to cross the junction and recombine because the electric field pushes free electrons back into the *n*-type silicon and pushes holes back into the *p*-type silicon. Consequently, the depletion zone stops growing. Typically this region is less than a micrometer wide. Because of the lack of free charge carriers in it,

the depletion zone acts as an electrical insulator.

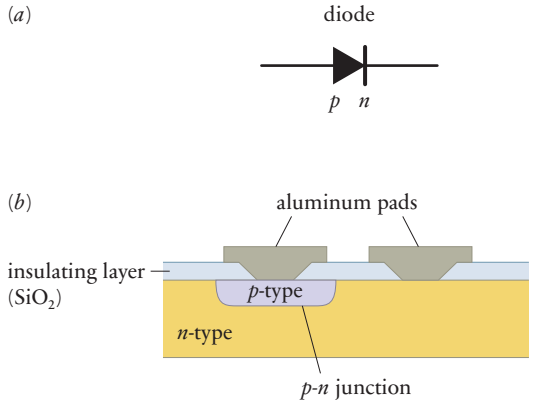
If we now connect the *n*-side of this diode to the positive terminal of a battery and the *p*-side to the negative terminal, the battery produces across the diode an electric field that points in the same direction as the electric field in the depletion zone (Figure 32.27*c*). The electric field of the battery pulls free electrons in the *n*-type silicon toward the positive terminal and pulls holes in the *p*-type silicon toward the negative terminal, broadening the (nonconducting) depletion zone. Connecting the battery in this manner therefore causes no flow of charge carriers in the diode.

When the battery is connected in the opposite direction, however, the depletion zone narrows as the battery's electric field pushes free electrons and holes toward the junction (Figure 32.27*d*). When the magnitude of the applied electric field created by the battery equals that of the electric field across the depletion zone, both types of free charge carriers can reach the junction, resulting in a current in the device carried both by free electrons and by holes.

As Figure 32.27 shows, a diode conducts current in one direction only: from the *p*-type side to the *n*-type side. The symbol for a diode is shown in **Figure 32.28*a***; the triangle points in the direction in which the diode conducts current (from the *p*-side to the *n*-side).

32.6 In the diode of Figure 32.28*a*, which way do holes travel? Which way do electrons travel?

Figure 32.28 (*a*) Circuit symbol for a diode. (*b*) Schematic of a diode made using integrated-circuit technology.



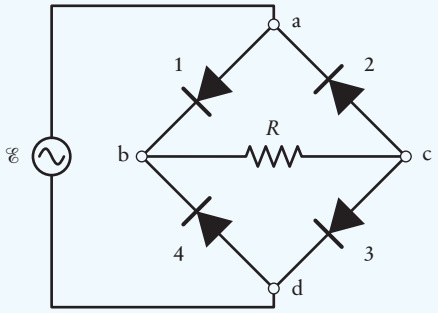
An *ideal diode* acts like a short circuit for current in the permitted direction and like an open circuit for current in the opposite direction. (That is not exactly how a diode behaves, but it's pretty close.)

32.7 Suppose a sinusoidally varying potential difference is applied across a diode connected in series with a resistor. Sketch the potential difference across the diode as a function of time, and then, on the same graph, sketch the current in the resistor as a function of time.

EXAMPLE 32.4 Rectifier

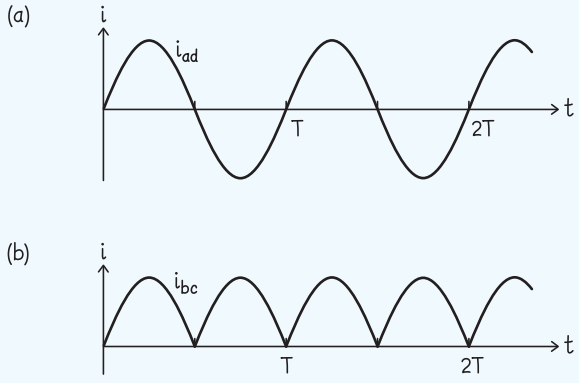
Consider the arrangement of ideal diodes shown in **Figure 32.29**. This arrangement, called a *rectifier*, converts alternating current (AC) to direct current (DC). Sketch a graph showing, for a sinusoidally alternating source, the current in the resistor in the direction from b to c as a function of time.

Figure 32.29 Example 32.4.



❶ **GETTING STARTED** Because the source is alternating, the current in the circuit periodically reverses direction. During part of the cycle the charge carriers creating the current flow clockwise through the source, and during another part of the cycle they flow counterclockwise. The diodes, however, conduct current in one direction only. I begin by making a sketch of the current between a and d, taking the direction from a to d to be positive (**Figure 32.30*a***).

Figure 32.30



❷ **DEVISE PLAN** In an ideal diode, the charge carriers can flow only in the direction in which the triangle in the diode symbol points. I shall determine which diodes allow charge

(Continued)

CONCEPTS

carriers to flow when the current direction is clockwise and when it is counterclockwise. I can then determine in each case which way the charge carriers flow through the resistor.

3 EXECUTE PLAN When the current in the circuit is clockwise, only diodes 1 and 3 are conducting, so the current direction is abcd. When the current in the circuit is counterclockwise ($i_{ad} < 0$), only diodes 2 and 4 are conducting, so the current direction is dbca. At all instants, the current in the resistor points in the same direction: from b to c. This means that i_{bc} is positive regardless of whether i_{ad} is positive or negative. Whenever i_{ad} is negative, the diodes reverse the direction of the current in the resistor, so i_{bc} is always positive and my graph is as shown in Figure 32.30*b*. ✔

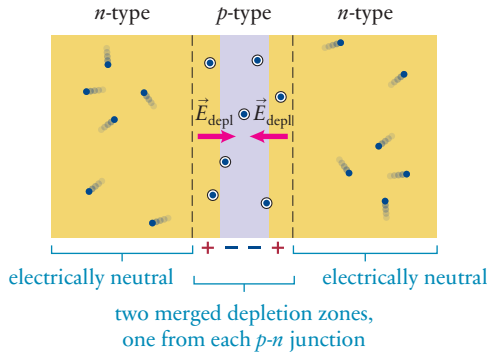
4 EVALUATE RESULT The arrangement of diodes keeps the current from b to c always in the same direction, even though the current from a to d alternates in direction. It makes sense, then, that this arrangement of diodes is called a *rectifier*.

Figure 32.28*b* shows how a diode may be constructed as part of an integrated circuit (a computer chip, for example). An aluminum pad (part of the metal wire connecting the diode to the rest of the circuit) is in contact with a small *p*-type region of silicon, which is surrounded by a larger *n*-type region that is in contact with a second aluminum pad. The *p*-*n* junction forms at the interface between the *p*- and *n*-type regions. A thin layer of silicon oxide (SiO_2) insulates the aluminum from the underlying silicon except where electrical contact is needed. On a modern computer chip, the entire device is only a few micrometers wide.

Another important circuit element in modern electronics is the **transistor,** a device that allows current control that is more precise than the on/off control of a diode. A transistor consists of a thin layer of one type of doped semiconductor sandwiched between two layers of the opposite type of doped semiconductor. **Figure 32.31**, for example, shows an *npn-type bipolar transistor*—a thin layer of *p*-type silicon sandwiched between two thicker regions of *n*-type silicon.* If the *p*-type layer is thin, the depletion zone formed at the left *p*-*n* junction merges with the depletion zone formed at the right *p*-*n* junction. The merged depletion zones form one wide depletion zone.

When a potential difference is applied across such a transistor (**Figure 32.32*a***), the depletion zone across junction 1 disappears, but that across junction 2 grows, shifting the depleted region toward the positive terminal of the battery. While charge carriers can now cross junction 1 where the depletion zone has disappeared, the (shifted) depletion zone that still exists prohibits their movement, which means no current in the transistor. For historical reasons, the *n*-type region connected to the negative terminal is called the *emitter*, the *n*-type region connected to the positive terminal is called the *collector*, and the *p*-type layer is called the *base*. If the direction of the applied potential difference is reversed, the roles of the emitter and the collector are also reversed, and there is still no current in the transistor.

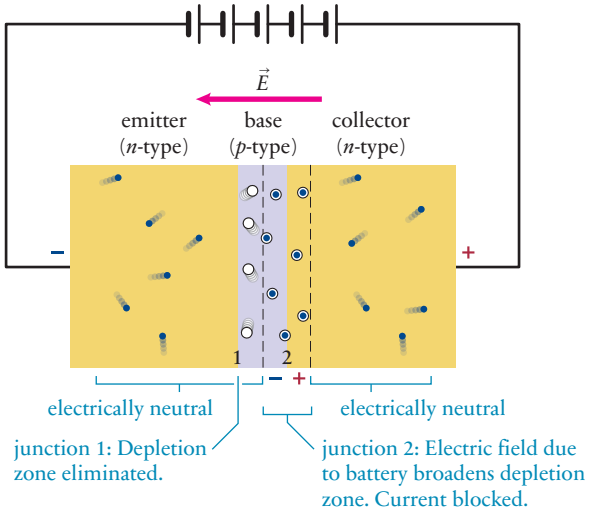
Figure 32.31 Schematic of an *npn*-type bipolar transistor, showing charge distribution and depletion zones for both *p*-*n* junctions.



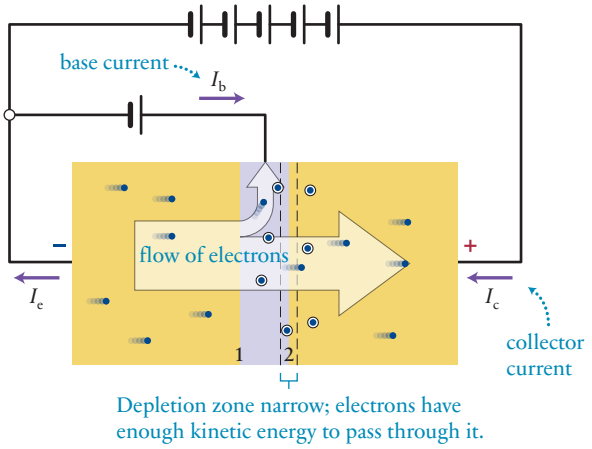
*Transistors in which a thin layer of *n*-type silicon is sandwiched between pieces of *p*-type silicon, called *pnp*-type bipolar transistors, are also used.

Figure 32.32 How an *npn*-type bipolar transistor works.

(*a*) Potential difference applied from collector to emitter only



(*b*) Potential difference also applied from base to emitter



CONCEPTS

Figure 32.33 Circuit symbol for an *npn*-type bipolar transistor.

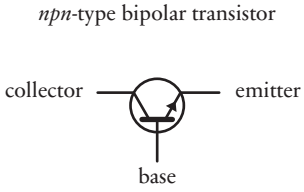
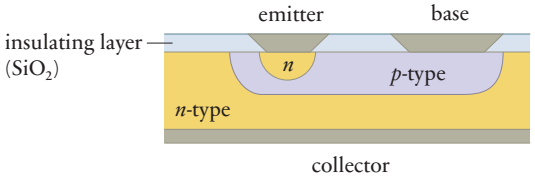


Figure 32.35 Schematic of an *npn*-type bipolar transistor made using integrated-circuit technology.



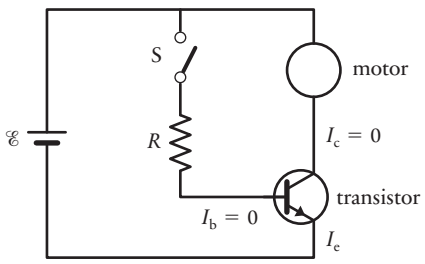
The situation changes drastically when, in addition to the potential difference between the emitter and the collector, a small potential difference is applied between the emitter and the base (Figure 32.32*b*). Adding this potential difference, called a *bias* or *bias potential difference*, makes the depletion zone much thinner than it is in Figure 32.32*a* because the formerly negatively charged region of this zone is brought to a positive potential, restoring mobile holes to that region. Because the emitter-base junction is conducting (remember, the depletion zone at junction 1 has disappeared), electrons now start flowing from the emitter toward the base. Once in the base, three things happen: (1) a small fraction of the electrons recombine with holes in the base, (2) electrons are attracted by the positive charge on the collector and have sufficient kinetic energy to pass straight through the very thin depletion zone, producing a collector current I_c, and (3) electrons diffuse through the base toward the positively charged end of the base, causing a small base current I_b. In a typical bipolar transistor, the collector current is 10 to 1000 times greater than the base current.

The circuit symbol for an *npn*-type bipolar transistor is shown in **Figure 32.33**.

Transistors are ubiquitous in modern electronics. In most applications, the transistor functions as either a switch or a current amplifier. If we consider I_b to be the input current and I_c the output current, the transistor acts as a switch in which I_b turns on and controls I_c. As a current amplifier, a small current I_b produces a much larger current I_c.

For electrical devices that draw large currents, it is useful to switch the device on and off with a mechanical switch wired in parallel with the device, rather than in series, so that the current in the device does not have to pass through the switch. **Figure 32.34** shows a circuit that utilizes such switching. When switch S is open, the base current is zero, and so the collector current (and therefore the current in the device) is zero. When switch S is closed, the small current from base to emitter causes a large current from collector to emitter that turns on the motor.

Figure 32.34 Circuit in which a bipolar transistor is used to turn a motor on and off.



32.8 In a bipolar transistor, what relationship, if any, exists among I_b, I_c, and the emitter current I_e?

Figure 32.35 shows how an *npn*-type bipolar transistor can be fabricated. A drawback of this type of transistor, however, is that a continuous small current through the base is required to make the transistor conducting. For this reason, another type of transistor, called the *field-effect transistor*, is used much more frequently. **Figure 32.36*a*** on the next page shows the configuration of one. Two *n*-type wells are made in a piece of *p*-type material. The *p*-type material between the two wells is covered with a nonconducting oxide layer (typically SiO_2) and then with a metal layer called the *gate*. The two *n*-type wells are called the *source* and the *drain* (the *n*-type well that is kept at a higher potential is the drain).

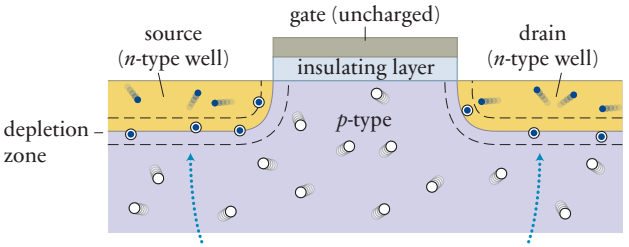
Because of the depletion zones between the *p*-type and *n*-type materials, no charge carriers can flow from the source to the drain (or vice versa). The nonconducting layer between the gate and the *p*-type material prevents charge carriers from traveling between the gate and the rest of the device.

If the gate is given a positive charge, as in Figure 32.36*b*, the (positively charged) holes just underneath the gate are pushed away, forming underneath the gate an additional depletion zone that connects the depletion zones around the two *n*-*p* junctions. If the positive charge is made large enough, electrons from the source and from the drain are pulled underneath the gate, forming an *n*-type channel below the gate (Figure 32.36*c*). This channel allows charge carriers to flow between the source and the drain. The gate thus controls the current between the source and the drain, just as the base in an *npn*-type bipolar transistor controls the current between the emitter and the collector. (The difference is that there is no current in the gate in a field-effect transistor.) Applying a positive charge to the gate is often referred to as putting a positive *bias* on the gate.

CONCEPTS

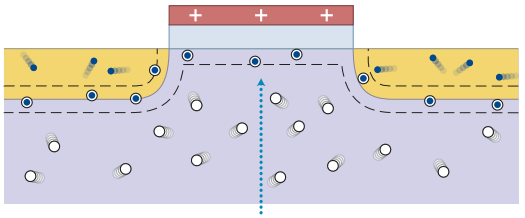
Figure 32.36 How a field-effect transistor works.

(*a*) Field-effect transistor with uncharged gate



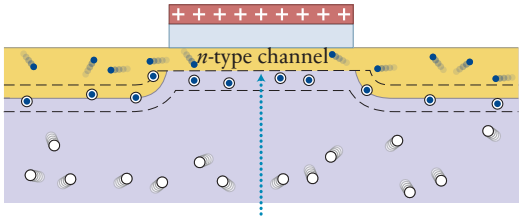
Uncharged gate: Separate depletion zones at *p*-*n* junctions.

(*b*) Small positive charge on gate attracts electrons to gate and extends depletion zone below gate



Small gate charge causes depletion zone to extend beneath gate.

(*c*) Large positive charge on gate attracts more electrons to gate and causes *n*-type channel, which connects source and drain



Strong gate charge pushes depletion zone away; conducting *n*-type channel now connnects source and drain.

Figure 32.37 (*a*) Circuit symbol for a field-effect transistor. (*b*) Schematic of a field-effect transistor made using integrated-circuit technology.

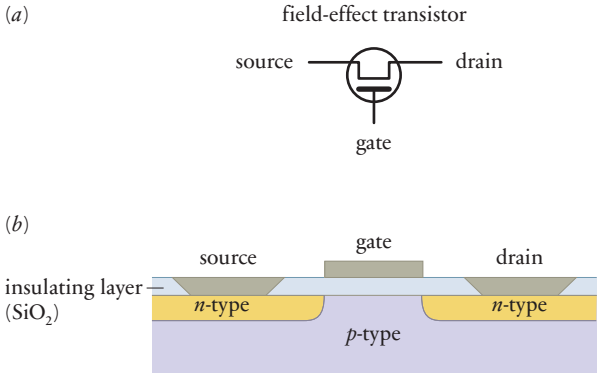
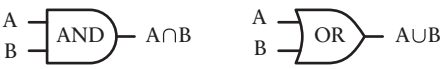


Figure 32.37*a* shows the circuit symbol for a field-effect transistor, and Figure 32.37*b* shows how this type of transistor can be realized in an integrated circuit. This type of transistor has two advantages over the bipolar transistor shown in Figure 32.35. First, all the terminals in the field-effect transistor are on the same side of the chip, making fabrication in integrated circuits much easier. Second, the current between the source and the drain is controlled by the charge on the gate, allowing a potential difference rather than a current to be used to control the source-drain current. Because no current is leaving the gate, no energy is required to keep current flowing from the source to the drain.

Field-effect transistors are widely used in devices called *logic gates*, which are the building blocks of computer processors and memory. A logic gate takes two input signals and provides an output after performing a logic operation on the input signals. For example, in a so-called AND gate, the output potential is nonzero with respect to ground only when both inputs are at positive potential with respect to ground. In an OR gate, the output potential is nonzero when either input potential is positive. The symbols used for these gates in circuit diagrams are shown in **Figure 32.38**; the inputs are on the left, and the output is on the right. In analyzing these circuits, we'll make the simplifying assumption that a transistor is just a switch that is open (off) when the potential of the gate is either at ground or negative with respect to ground and is closed (on) when the gate is at a positive potential.

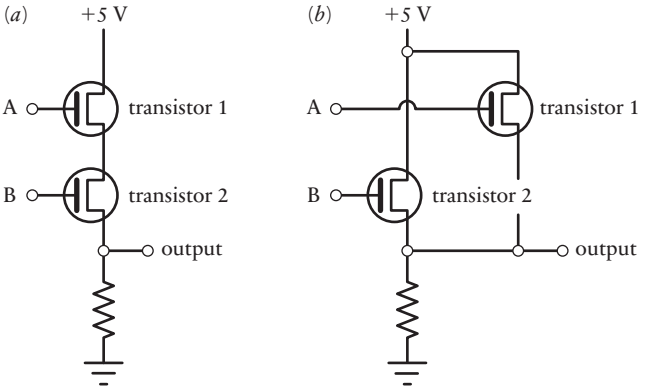
Figure 32.38 Circuit symbols for AND and OR logic gates.



CONCEPTS

32.9 Circuit diagrams for two logic gates are shown in **Figure 32.39**. Which is the AND gate, and which is the OR gate? Explain briefly how each one works.

Figure 32.39 Checkpoint 32.9.



Self-quiz

1. At the instant shown in **Figure 32.40**, the potential difference across the capacitor is half its maximum value and the charge on the plates is increasing. Draw the direction of the current and sketch the magnetic field at this instant. Is the magnitude of current increasing or decreasing?
2. Construct a phasor diagram representing the current and potential difference in Figure 32.10 at $t = T/4$, $T/2$, and $3T/4$.
3. **Figure 32.41** shows the time-varying potential difference and current for the circuit of Figure 32.8. At the instant labeled t_a, what are the charge on the capacitor plates and the direction of the current?

Figure 32.40

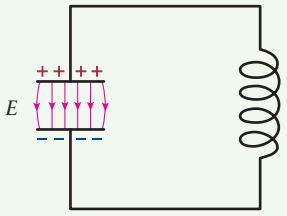
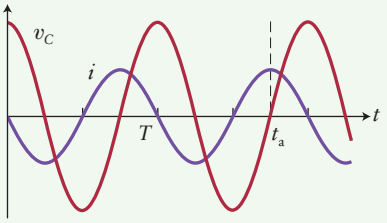
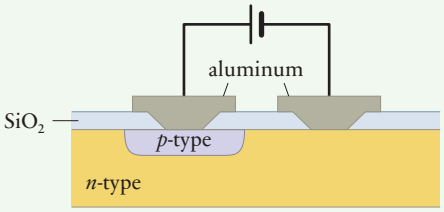


Figure 32.41



4. Is there any current in a diode connected as shown in **Figure 32.42**?

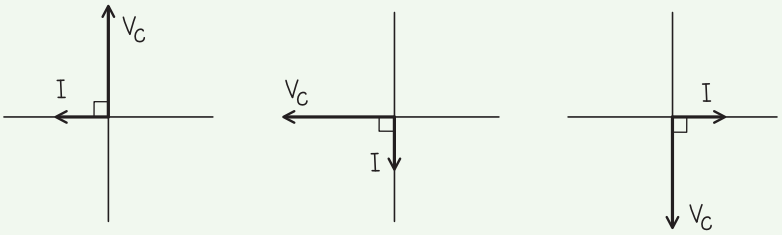
Figure 32.42



Answers

1. Your sketch should show the current directed counterclockwise. The magnetic field in the center of the coil points up the page according to the right-hand dipole rule (assuming we are looking down on the top of the coil in Figure 32.40). Because the current is zero when the capacitor has maximum charge, the magnitude of the current is decreasing at the instant shown in Figure 32.40.
2. See **Figure 32.43**. At $t = T/4$ the potential difference phasor V_C points along the positive y axis because the potential difference reaches its maximum positive value at this instant, and the current phasor I points along the negative x axis because it leads the current by 90°. Each quarter cycle both phasors rotate 90° counterclockwise.

Figure 32.43



3. Because the potential difference across the capacitor is zero at instant t_a, the charge on the plates must be zero. The current is a maximum at this instant and is directed clockwise.
4. Yes. The holes in the p-type material move away from the positive terminal, and the electrons move toward it. According to Figure 32.27*d*, this flow shrinks the depletion zone, the charge carriers can flow, and so there is a current.

32.5 Reactance

Let us now develop a mathematical framework for analyzing alternating-current circuits. The instantaneous emf supplied by an AC source is customarily written as

$$\mathscr{E} = \mathscr{E}_{\max} \sin \omega t, \tag{32.1}$$

where $\mathscr{E}_{\max}$ is the maximum value of the emf, typically called the *peak value* or *amplitude* (see Section 15.1), $\omega = 2\pi f$ is the angular frequency of oscillation in inverse seconds (Section 15.5), and f is the frequency in hertz. Most generators have frequencies of 50 Hz or 60 Hz. Audio circuits typically operate at kilohertz frequencies, radio transmitters at 10^8 Hz, for instance, and computer chips at 10^9 Hz. It's very important to remember to convert frequencies in hertz (cycles per second) to angular frequencies in s^{-1} when ω appears in the equations.

Note that the initial phase for the emf as written in Eq. 32.1 is zero. When we make this choice, the source emf serves as the reference for phase in the circuit.

Let's begin by revisiting the circuit from Exercise 32.1—a resistor connected to an AC source (**Figure 32.44**). At any instant, Ohm's law relates the potential difference across the resistor to the current in it, just as it does for DC circuits:

$$v_R = iR. \tag{32.2}$$

The only difference between Eq. 32.2 and Ohm's law for DC circuits (Eq. 31.11) is that the potential difference and the current in Eq. 32.2 oscillate in time.

Applying the loop rule to this circuit gives the AC version of Eq. 31.23:

$$\mathscr{E} - iR = 0. \tag{32.3}$$

Equations 32.2 and 32.3 show that the potential difference across the load equals the emf supplied by the source (as we would expect):

$$v_R = \mathscr{E} = \mathscr{E}_{\max} \sin \omega t. \tag{32.4}$$

Figure 32.44 AC circuit consisting of a resistor connected across the terminals of an AC source.

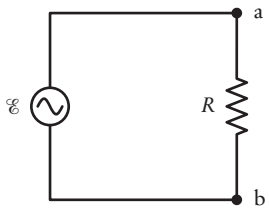
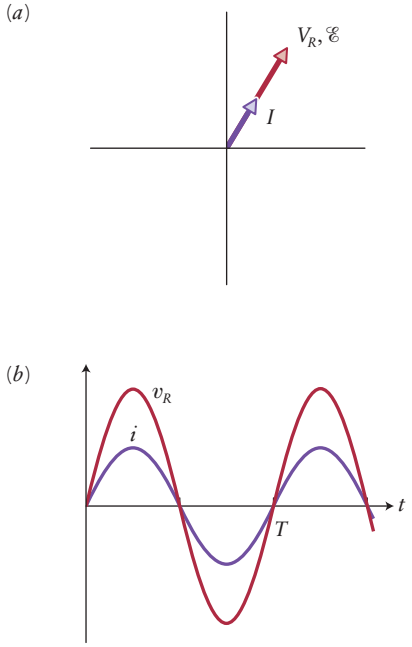


Figure 32.45 (*a*) Phasor diagram and (*b*) graph showing time dependence of i and v_R for the circuit shown in Figure 32.44.



32.10 (*a*) In Figure 32.44, is the potential at point a higher or lower than the potential at b when the current direction is clockwise through the circuit? (*b*) If we define such a current to be positive, is $\mathscr{E}$ positive or negative? Express v_R in terms of the potential at a and the potential at b. (*c*) Half a cycle later, when the current is negative, is $\mathscr{E}$ positive or negative? Express v_R again in terms of the potential at a and the potential at b.

Using Eqs. 32.2 and 32.4, we can write the current in the resistor as

$$i = \frac{v_R}{R} = \frac{\mathscr{E}_{\max} \sin \omega t}{R} = I \sin \omega t, \tag{32.5}$$

where $I = \mathscr{E}_{\max}/R$ is the amplitude of the current. Note that the current and the potential difference both oscillate at angular frequency ω and are in phase, as we concluded in Exercise 32.1. If we write $v_R = V_R \sin \omega t$, we see that the amplitudes of the current and the potential difference satisfy the relationship

$$V_R = IR. \tag{32.6}$$

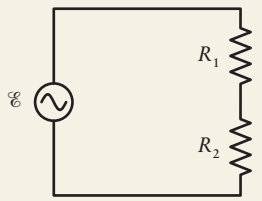
Figure 32.45 shows the corresponding phasor diagram and time dependence of v_R and i.

QUANTITATIVE TOOLS

EXERCISE 32.5 AC circuit with two resistors

In **Figure 32.46**, the resistances are $R_1 = 100\ \Omega$ and $R_2 = 60\ \Omega$, the amplitude of the emf is $\mathscr{E}_{max} = 160$ V, and its frequency is 60 Hz. (*a*) What is the amplitude of the potential difference across each resistor? (*b*) What is the instantaneous potential difference across each resistor at $t = 50$ ms?

Figure 32.46 Exercise 32.5.



SOLUTION I analyze this circuit just as I would analyze a DC circuit containing two resistors, except now I must keep in mind that the current and potential differences are oscillating. The resistance of the load is

$$R_{load} = R_1 + R_2,$$

and the instantaneous current in the load is

$$i = \frac{\mathscr{E}}{R_{load}} = \frac{\mathscr{E}}{R_1 + R_2}.$$

(*a*) Because the current and the emf are in phase, they reach their maximum values at the same instant. As a result, the amplitude (maximum value) of the current is given by the amplitude of the emf divided by the resistance:

$$I = \frac{\mathscr{E}_{max}}{R_1 + R_2} = 1.0\ \text{A}.$$

The potential differences across the resistors are in phase with the current, and so I calculate the amplitude of the potential differences from the amplitude of the current using Eq. 32.6:

$$V_{R_1} = IR_1 = (1.0\ \text{A})(100\ \Omega) = 100\ \text{V}$$

$$V_{R_2} = IR_2 = (1.0\ \text{A})(60\ \Omega) = 60\ \text{V}. \checkmark$$

(*b*) I can use Eq. 32.5 to calculate the instantaneous value of the current:

$$i = (1.0\ \text{A})\sin(2\pi \cdot 60\ \text{Hz} \cdot 0.050\ \text{s}) = 0.$$

(In 50 ms, three full cycles at 60 Hz take place.) Because the current is zero at 50 ms, the potential differences v_{R_1} and v_{R_2} at 50 ms are also zero. ✔

Next consider a capacitor connected to an AC source (**Figure 32.47**). Because the capacitor and the source are connected to each other, we have

$$\mathscr{E} = v_C, \tag{32.7}$$

and so the potential difference across the capacitor is

$$v_C = \mathscr{E}_{max} \sin \omega t = V_C \sin \omega t. \tag{32.8}$$

At any instant the potential difference across the capacitor and the charge on the upper plate are related by (see Eq. 26.1)

$$\frac{q}{v_C} = C, \tag{32.9}$$

where the potential difference v_C and the charge q on the plate oscillate in time. The charge on the upper capacitor plate is thus

$$q = Cv_C = CV_C \sin \omega t, \tag{32.10}$$

and the current is the rate of change of the charge on the plate:

$$i = \frac{dq}{dt} = \frac{d}{dt}(CV_C \sin \omega t) = \omega CV_C \cos \omega t. \tag{32.11}$$

Using the identity $\cos \alpha = \sin\left(\alpha + \frac{\pi}{2}\right)$, we can rewrite this as

$$i = \omega CV_C \sin\left(\omega t + \frac{\pi}{2}\right) = I \sin\left(\omega t + \frac{\pi}{2}\right). \tag{32.12}$$

We now see that v_C and i are not in phase: i reaches its maximum value one-quarter period before v_C reaches its maximum value (**Figure 32.48** on the next page), as we found in Section 32.2.

Figure 32.47 AC circuit consisting of a capacitor connected across the terminals of an AC source.

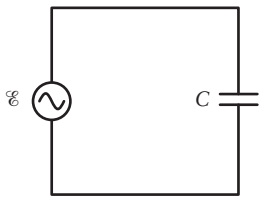
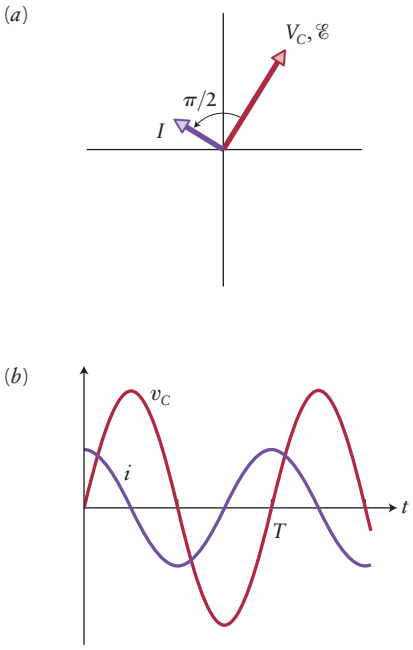


Figure 32.48 (*a*) Phasor diagram and (*b*) graph showing time dependence of i and v_C for the circuit of Figure 32.47. The phasor diagram shows the relative phase of i and v_C.



The current in the capacitor of Figure 32.47 is not simply proportional to the potential difference across the capacitor because the two are out of phase. However, the *amplitude* of the current is proportional to the amplitude of the potential difference: $I = \omega C V_C$. Rewriting this to express V_C in terms of I gives

$$V_C = \frac{I}{\omega C}. \tag{32.13}$$

Note how this expression differs from the expression for a circuit that consists of only an AC source and a resistor, $V_R = IR$ (Eq. 32.6), where R is the proportionality constant between V and I. In Eq. 32.13, the proportionality constant is no longer a resistance (though it still has units of ohms). In circuits that contain capacitors and/or inductors, we use the general name **reactance** for the proportionality constant between the potential difference amplitude and the current amplitude. From Eq. 32.13 we see that this proportionality constant for a circuit that contains a capacitor is $1/\omega C$, and we call this constant the *capacitive reactance* X_C:

$$X_C \equiv \frac{1}{\omega C}, \tag{32.14}$$

so Eq. 32.13 becomes

$$V_C = IX_C. \tag{32.15}$$

Reactance is a measure of the opposition of a circuit element to a change in current. Unlike resistance, reactance is frequency dependent. At low frequency, the capacitive reactance X_C is large, which means that the amplitude of the current is small for a given value of V_C. At zero frequency, the current $I = \omega C V_C$ is zero, as it should be. (There is no direct current in a capacitor because the capacitor is just like an open circuit!) The higher the frequency of the source, the smaller the capacitive reactance and the greater the current (the less the capacitor opposes the alternating current).

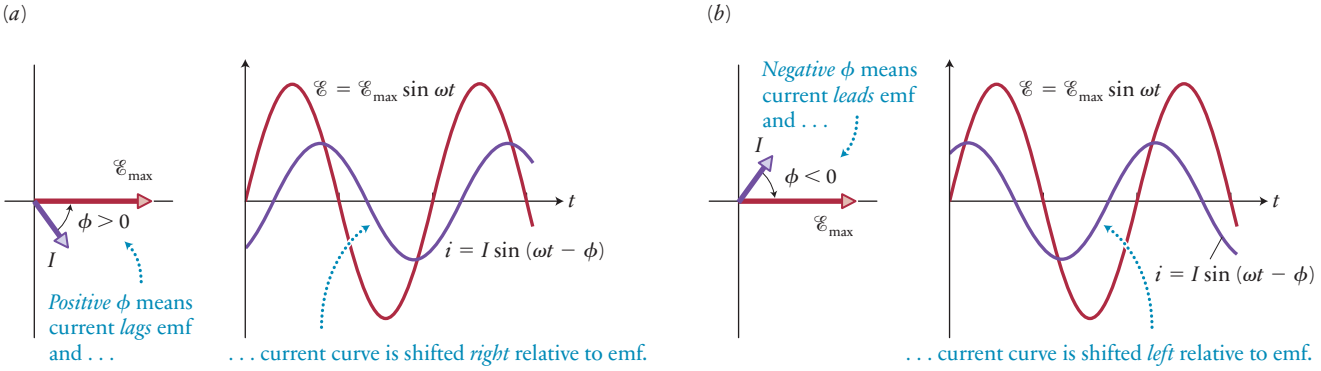
Often, when analyzing AC circuits, the only things we are interested in are the amplitudes of the currents and potential differences. The capacitive reactance allows us to calculate the amplitude of the current in the capacitor directly from the amplitude of the potential difference across it—in this case, the emf of the source.

It is conventional to write the current in an AC circuit in the form

$$i = I \sin(\omega t - \phi), \tag{32.16}$$

where ϕ is called the **phase constant.** The negative sign in front of the phase constant is chosen so that a positive ϕ corresponds to shifting the curve for the current to the right, in the positive direction along the time axis, and a negative ϕ corresponds to shifting the curve to the left, in the negative direction along this axis (**Figure 32.49**).

Figure 32.49 Positive and negative phase constant.



QUANTITATIVE TOOLS

The phase constant represents the phase difference between the source emf and the current. It is measured from the current phasor to the source emf phasor with the counterclockwise direction being positive (Figure 32.49). As a result, when the current leads the source emf, ϕ is negative; when the current lags the source emf, ϕ is positive.

Comparing Eqs. 32.12 and 32.16, we see that for the capacitor-AC source circuit of Figure 32.47, $\phi = -\pi/2$, as shown in Figure 32.48*a*. The negative phase constant means that the current leads the source emf. The curve for i is shifted to the left relative to the curve for v_C, as shown in Figure 32.48*b*. As you can see from the figure, when the capacitor has maximum charge (v_C maximum), the current is zero because at that instant the current reverses direction as the capacitor begins discharging. The current reaches its maximum value when the capacitor is completely discharged ($v_C = 0$).

32.11 As in the *LC* circuit discussed in Section 32.1, the current in the circuit of Figure 32.47 oscillates. If we think of v_C as corresponding to the position of the simple harmonic oscillator described in Section 15.5, what property of the circuit of Figure 32.47 corresponds to the velocity of the oscillator?

Finally, consider an inductor connected to an AC source (**Figure 32.50**). Because the inductor and the source are connected to each other, we have

$$\mathscr{E} = v_L, \tag{32.17}$$

so the potential difference across the inductor is

$$v_L = \mathscr{E}_{\max} \sin \omega t = V_L \sin \omega t. \tag{32.18}$$

In Chapter 29 we saw that a changing current in an inductor causes an induced emf (Eq. 29.19):

$$\mathscr{E}_{\text{ind}} = -L\frac{di}{dt}. \tag{32.19}$$

The negative sign in this expression means that the potential decreases across the inductor in the direction of increasing current. Consequently, in Figure 32.50, the potential at b is lower than the potential at a when the current is increasing clockwise around the circuit. However, for consistency with Eq. 32.3, we always measure the potential difference v_L from b to a, just as we did with the AC source-resistor circuit of Figure 32.44. Therefore the sign of the potential difference across the inductor is the opposite of the sign in Eq. 32.19:

$$v_L = L\frac{di}{dt}. \tag{32.20}$$

We obtain the current in the circuit by substituting Eq. 32.18 into Eq. 32.20:

$$L\frac{di}{dt} = V_L \sin \omega t \tag{32.21}$$

$$di = \frac{V_L}{L} \sin \omega t \, dt. \tag{32.22}$$

To obtain the current, we integrate this expression:

$$i = \frac{V_L}{L}\int \sin \omega t \, dt = -\frac{V_L}{\omega L} \cos \omega t. \tag{32.23}$$

Figure 32.50 AC circuit consisting of an inductor connected across the terminals of an AC source.

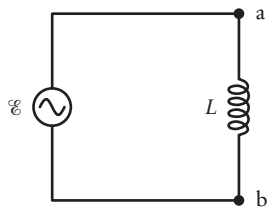
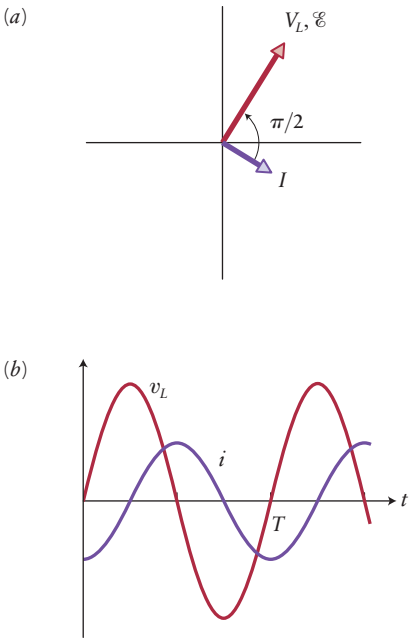


Figure 32.51 (*a*) Phasor diagram and (*b*) graph showing time dependence of i and v_L for the circuit of Figure 32.50. The phasor diagram shows the phase difference $\phi = \pi/2$ between i and v_L.



The amplitude of the current is thus

$$I = \frac{V_L}{\omega L}, \tag{32.24}$$

and using the identity $\cos \omega t = -\sin\left(\omega t - \frac{\pi}{2}\right)$, we get

$$i = I \sin\left(\omega t - \frac{\pi}{2}\right). \tag{32.25}$$

The phase constant is $\phi = +\pi/2$, which means the current lags the source by 90°, as shown in **Figure 32.51**.

Just as we defined a capacitive reactance for a circuit that contains a capacitor, we define the *inductive reactance* X_L for a circuit that contains an inductor as the constant of proportionality between the amplitudes V_L and I in the circuit. From Eq. 32.24 we see that this proportionality constant is ωL:

$$X_L \equiv \omega L, \tag{32.26}$$

so that

$$V_L = IX_L. \tag{32.27}$$

Inductive reactance, like capacitive reactance, has units of ohms and depends on the frequency of the AC source. However, X_L increases with increasing frequency, so, at a given potential difference, the amplitude of the current is greatest at zero frequency and decreases as the frequency increases. This makes sense because for a constant current, an inductor is just a conducting wire and does not impede the current; as the frequency of the AC source increases, the emf induced across the inductor increases.

EXAMPLE 32.6 Oscillating inductor

When a 3.0-H inductor is the only element in a circuit connected to a 60-Hz AC source that is delivering a maximum emf of 160 V, the current amplitude is I. When a capacitor is the only element in a circuit connected to the same source, what must the capacitance be in order to have the current amplitude again be I?

❶ **GETTING STARTED** I begin by identifying the information given in the problem statement: $\mathscr{E}_{max} = 160$ V, angular frequency $\omega = 2\pi(60\text{ Hz})$, and inductance $L = 3.0$ H. The problem asks me to compare two circuits, one with an inductor connected to an AC source and the other with a capacitor connected to the same source. What I must determine is the capacitance value that makes the current amplitude the same in the two circuits.

❷ **DEVISE PLAN** For both circuits the potential difference across the load equals the source emf, so $\mathscr{E}_{max} = V_C = V_L$. I can use Eqs. 32.26 and 32.27 to get an expression for V_L in terms of I, from which I can express I in the inductor circuit in terms of V_L, ω, and L. Next I can use Eqs. 32.14 and 32.15 to get an expression for V_C in terms of I. I can then substitute this into my first expression for I and obtain an expression for C that contains only known quantities.

❸ **EXECUTE PLAN** Substituting the inductive reactance from Eq. 32.26, $X_L = \omega L$, into Eq. 32.27, $V_L = IX_L$, I get $V_L = I\omega L$, so the amplitude of the current is

$$I = \frac{V_L}{\omega L}. \tag{1}$$

Substituting the capacitive reactance from Eq. 32.14, $X_C = 1/\omega C$, into Eq. 32.15, $V_C = IX_C$, I get

$$V_C = \frac{I}{\omega C}. \tag{2}$$

Solving Eq. 2 for C and substituting Eq. 1 for I give

$$C = \frac{I}{\omega V_C} = \frac{V_L}{\omega^2 V_C L} = \frac{1}{\omega^2 L}$$

$$= \frac{1}{(2\pi \cdot 60\text{ Hz})^2(3.0\text{ H})} = 2.3 \times 10^{-6}\text{ F}. \checkmark$$

❹ **EVALUATE RESULT** To check my answer, I can calculate the inductive and capacitive reactances from Eqs. 32.26 and 32.14, respectively: $X_L = \omega L = (2\pi \cdot 60\text{ Hz})(3.0\text{ H}) = 1.1\text{ k}\Omega$ and $X_C = 1/\omega C = 1/(2\pi \cdot 60\text{ Hz})(2.3 \times 10^{-6}\text{ F}) = 1.1\text{ k}\Omega$. The two reactances are identical, as I expect given that they yield the same current amplitude for the same AC source.

QUANTITATIVE TOOLS

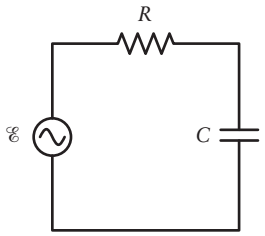
32.12 For the three circuits discussed in this section (AC source with resistor, capacitor, or inductor), sketch for a given emf amplitude (*a*) the resistance or reactance as a function of angular frequency ω and (*b*) the current amplitude in the circuit as a function of ω. Explain the meaning of each curve on your graphs.

32.6 *RC* and *RLC* series circuits

When an AC source is connected to multiple circuit elements, either in series or in parallel, applying the loop rule becomes more complicated than for DC circuits because we need to add several oscillating potential differences that may be out of phase with one another. For example, suppose we have a resistor and a capacitor in series with an AC source (**Figure 32.52**), known as an *RC series circuit*. The loop rule states that

$$\mathscr{E} = v_R + v_C. \tag{32.28}$$

Figure 32.52 An *RC* series circuit, consisting of a resistor and a capacitor in series across the terminals of an AC source.



To compute the sum on the right side of this equation, we must add potential differences that vary sinusoidally at the same angular frequency ω but are out of phase. The combined potential difference v of two potential differences v_1 and v_2 that oscillate at the same angular frequency is

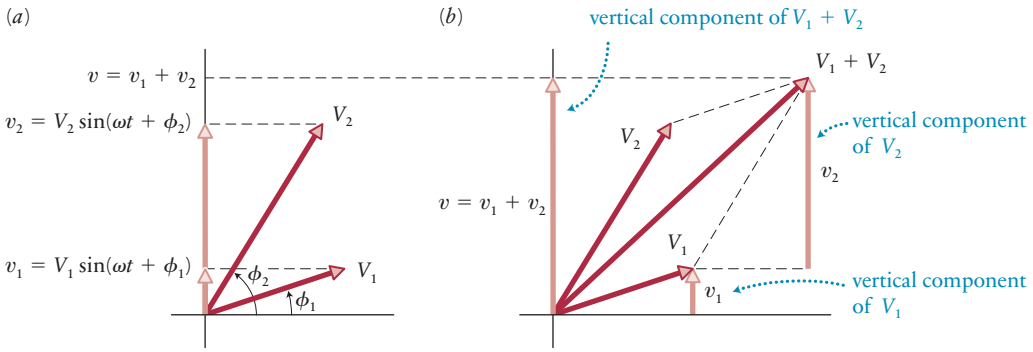
$$v = V_1 \sin(\omega t + \phi_1) + V_2 \sin(\omega t + \phi_2), \tag{32.29}$$

where ϕ_1 and ϕ_2 are the initial phases of the two potential differences. Calculating this sum algebraically gets very messy, but using phasors to calculate it simplifies things greatly.

Figure 32.53*a* shows the phasors that correspond to the two terms on the right in Eq. 32.29. Recall that the instantaneous value of the quantity represented by a rotating phasor equals the vertical component of the phasor (see Figure 32.11). Therefore, v at any instant equals the sum of the vertical components of the phasors that represent v_1 and v_2. This sum is equal to the vertical component of the vector sum $V_1 + V_2$ of the phasors, as shown in Figure 32.53*b*.

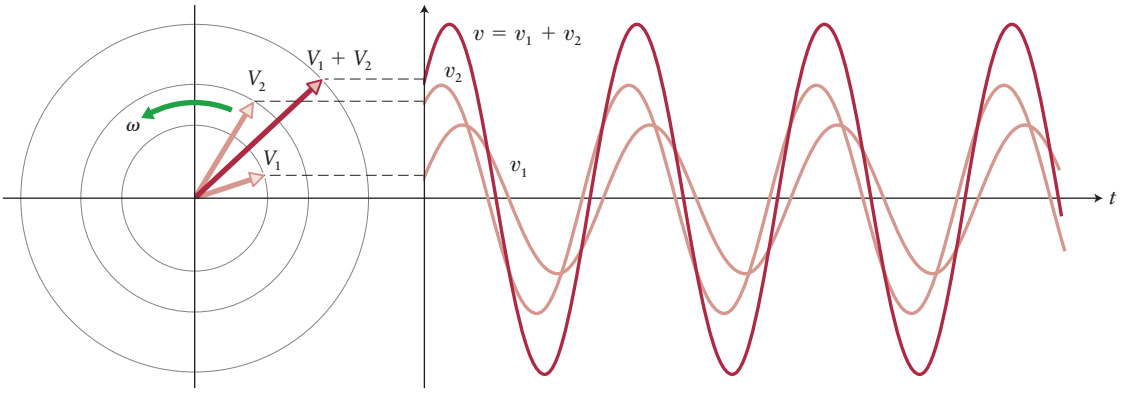
Note that the combined potential difference v oscillates at the same angular frequency as the two potential differences v_1 and v_2. Consequently, the three phasors V_1, V_2, and $V_1 + V_2$ rotate as a unit at angular frequency ω, as

Figure 32.53 (*a*) Phasor diagram for a system of two oscillating potential differences v_1 and v_2. (*b*) Vector diagram indicating that the vertical component of the vector sum of the phasors equals the sum of the vertical components of the individual phasors.



QUANTITATIVE TOOLS

Figure 32.54 Phasor diagram and graph showing time dependence of v_1, v_2, and $v = v_1 + v_2$ from Figure 32.53. All three phasors rotate as a unit at angular frequency ω.



shown in **Figure 32.54**. The phase relationship among the three phasors is constant, as is the phase relationship among the potential differences.

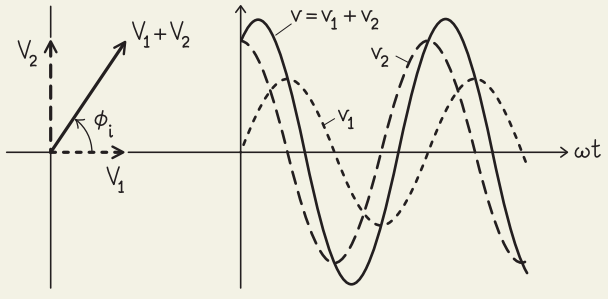
The next example shows how to apply these principles to a specific situation. To convince yourself that the phasor method is worthwhile, try adding the two original trigonometric functions algebraically after solving the problem using phasors!

EXAMPLE 32.7 Adding phasors

Use phasors to determine the sum of the oscillating potential differences $v_1 = (2.0\ \text{V}) \sin \omega t$ and $v_2 = (3.0\ \text{V}) \cos \omega t$.

❶ **GETTING STARTED** I begin by making a graph showing the time dependence of v_1 and v_2, and I draw the corresponding phasors V_1 and V_2 to the left of my graph (**Figure 32.55**). I add to my phasor diagram the phasor $V_1 + V_2$, which is the phasor that represents the potential difference sum $v_1 + v_2$ that I must determine. Using phasor $V_1 + V_2$, I can sketch the time dependence of the sum $v_1 + v_2$ by tracing out the projection of phasor $V_1 + V_2$ onto the vertical axis of my $V(\omega t)$ graph as this phasor rotates counterclockwise from the starting position I drew.

Figure 32.55



❷ **DEVISE PLAN** To obtain an algebraic expression for $v_1 + v_2$, I first write the oscillating potential differences in the form $v_1 = V_1 \sin(\omega t + \phi_1)$ and $v_2 = V_2 \sin(\omega t + \phi_2)$. Comparing these expressions with the given potential differences, I see that $V_1 = 2.0$ V, $\phi_1 = 0$, and $V_2 = 3.0$ V. In order to determine ϕ_2, I use the trigonometric identity $\cos(\omega t) = \sin(\omega t + \pi/2)$, and so my given information $v_2 = (3.0\ \text{V}) \cos \omega t = (3.0\ \text{V}) \sin(\omega t + \pi/2)$ tells me that $\phi_2 = \pi/2$. The sum $v_1 + v_2$ is a sinusoidally varying function that can be written as $v_1 + v_2 = A \sin(\omega t + \phi_i)$. The amplitude A is equal to the length of the phasor $V_1 + V_2$, and from my sketch I see that the initial phase ϕ_i is given by the angle between $V_1 + V_2$ and V_1.

❸ **EXECUTE PLAN** The length of the phasor $V_1 + V_2$ is given by the Pythagorean theorem applied to the right triangle containing ϕ_i in my phasor diagram:

$$A = \sqrt{V_1^2 + V_2^2} = \sqrt{(3.0\ \text{V})^2 + (2.0\ \text{V})^2}$$
$$= \sqrt{13\ \text{V}^2} = 3.6\ \text{V}.$$

The tangent of the angle between $V_1 + V_2$ and V_1 is then

$$\tan \phi_i = \frac{V_2}{V_1} = \frac{3.0\ \text{V}}{2.0\ \text{V}} = 1.5,$$

so $\phi_i = \tan^{-1}(1.5) = 56°$.

Now that I have determined A and ϕ_i, I can write the sum of the two potential differences as

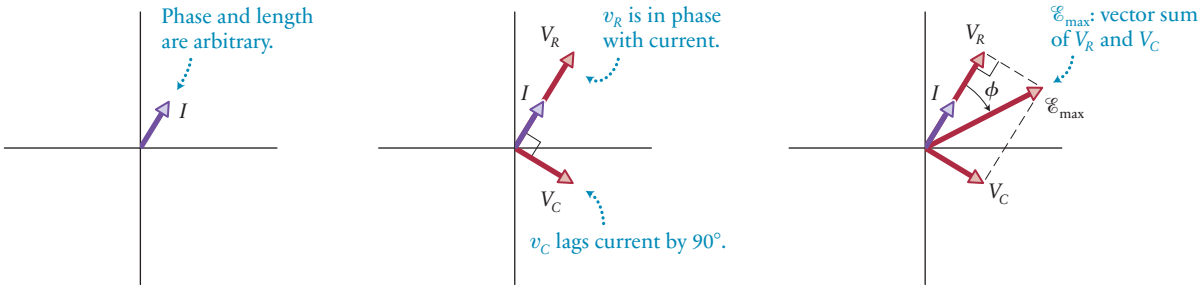
$$v_1 + v_2 = (3.6\ \text{V}) \sin(\omega t + 56°). \checkmark$$

❹ **EVALUATE RESULT** The amplitude of the sinusoidal function I obtained is 3.6 V, which is greater than the larger of the two phasors I added, as I expect. My answer shows that the sum of the two potential differences reaches its maximum when $\omega t + \phi_i = 90°$, or when $\omega t = 90° - \phi_i = 34°$. This conclusion agrees with my phasor diagram: The phasor $V_1 + V_2$ reaches the vertical position after it rotates through an angle of $90° - \phi_i = 90° - 56° = 34°$. (I could also verify my answer by adding the two original sine functions algebraically, but the trigonometry needed in that approach is tedious.)

QUANTITATIVE TOOLS

Figure 32.56 Steps involved in constructing a phasor diagram for the circuit in Figure 32.52. The diagram in part *c* indicates the phase of the current relative to the source emf.

(*a*) Draw current phasor (*b*) Add phasors for v_R and v_C (*c*) Add phasor for emf



32.13 Suppose you need to add two potential differences that are oscillating at *different* angular frequencies—say, 2 sin (ωt) and 3 cos ($2\omega t$). Can you use the phasor method described above to determine the sum? Why or why not?

Let us now return to the *RC* series circuit of Figure 32.52 and construct a phasor diagram in order to determine the amplitude and phase of the current in terms of the amplitude of the source emf and the resistance and capacitance of the circuit elements. From the current, we can calculate the potential differences across the circuit elements.

Because the circuit contains only one loop, the time-dependent current i is the same throughout. Therefore, we begin by drawing a phasor that represents i (**Figure 32.56*a***). We are free to choose the phase of this phasor because we have not yet specified the phase of any of the potential differences in the circuit. Also, the length we draw for phasor I is unimportant because it is the only current phasor for this circuit.

Next, we draw the phasors for v_R and v_C, the potential differences across the resistor and capacitor, respectively. We must get the relative phases right, and the lengths of the phasors must also be appropriately proportioned. Because the current is in phase with v_R (Figure 32.45*a*), we draw the corresponding phasor as shown in Figure 32.56*b*; its length is $V_R = IR$.

What about the phasor for v_C? We found previously that the current in a capacitor leads the potential difference across the capacitor by 90° (Figure 32.48*a*), which means we must draw the phasor for v_C 90° behind the phasor for i, as it is in Figure 32.56*b*. The length of this phasor is $V_C = IX_C$.

Finally, we need to draw the phasor for the emf supplied by the source. Phasor addition with the loop rule for this circuit (Eq. 32.28) tells us that the phasor $\mathscr{E}_{max}$ for the emf is the vector sum of the phasors V_R and V_C (Figure 32.56*c*). The amplitudes of the potential differences are related by

$$\mathscr{E}_{max}^2 = V_R^2 + V_C^2. \tag{32.30}$$

If we substitute $V_R = IR$ (Eq. 32.6) and $V_C = IX_C$ (Eq. 32.15), this becomes

$$\mathscr{E}_{max}^2 = (IR)^2 + (IX_C)^2 = I^2(R^2 + X_C^2) = I^2\left(R^2 + \frac{1}{\omega^2 C^2}\right). \tag{32.31}$$

Solving for I gives

$$I = \frac{\mathscr{E}_{max}}{\sqrt{R^2 + 1/\omega^2 C^2}}. \tag{32.32}$$

Remembering that $\mathscr{E}_{max} = V_{load}$, we see that even though this load includes both resistive and reactive elements, I is still proportional to V_{load}! The constant of proportionality is called the **impedance** of the load and is denoted by Z:

$$I = \frac{\mathscr{E}_{max}}{Z}. \tag{32.33}$$

The impedance of the load is a property of the entire load. It is measured in ohms and depends on the frequency for any load that contains reactive elements.

Impedance plays the same role in AC circuits that resistance plays in DC circuits. In fact, Eq. 32.33 can be thought of as the equivalent of Ohm's law for AC circuits. Equation 32.32 shows that, for an RC series circuit, Z depends on both R and C:

$$Z_{RC} \equiv \sqrt{R^2 + 1/\omega^2C^2} \quad (RC \text{ series combination}). \tag{32.34}$$

To express V_R and V_C in terms of $\mathscr{E}_{max}$, R, C, and ω, we use Eq. 32.32:

$$V_R = IR = \frac{\mathscr{E}_{max}R}{\sqrt{R^2 + 1/\omega^2C^2}} \tag{32.35}$$

$$V_C = IX_C = \frac{\mathscr{E}_{max}/\omega C}{\sqrt{R^2 + 1/\omega^2C^2}}. \tag{32.36}$$

To calculate the phase constant ϕ, the geometry shown in Figure 32.56*c* gives us, with Eqs. 32.6 and 32.15,

$$\tan\phi = -\frac{V_C}{V_R} = -\frac{IX_C}{IR} = -\frac{1}{\omega RC} \tag{32.37}$$

or

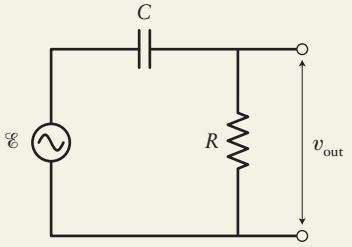
$$\phi = \tan^{-1}\left(-\frac{1}{\omega RC}\right) \quad (RC \text{ series circuit}). \tag{32.38}$$

The negative value of ϕ indicates that the current in an RC series circuit leads the emf, just as it does in an AC circuit with only a capacitor. As you can see in Figure 32.56*c*, however, the phase difference between the emf and the current in the RC series circuit is less than 90°.

EXAMPLE 32.8 High-pass filter

A circuit that allows emfs in one angular-frequency range to pass through essentially unchanged but prevents emfs in other angular-frequency ranges from passing through is called a *filter*. Such a circuit is useful in a variety of electronic devices, including audio electronics. An example of a filter, called a *high-pass filter*, is shown in Figure 32.57. Emfs that have angular frequencies above a certain angular frequency, called the *cutoff angular frequency* ω_c, pass through to the two output terminals marked v_{out}, but the filter attenuates the amplitudes of emfs that have frequencies below the cutoff value. (*a*) Determine an expression that gives, in terms of R and C, the cutoff angular frequency ω_c at which $V_R = V_C$. (*b*) Determine the potential difference amplitude v_{out} across the output terminals for $\omega \gg \omega_c$ and for $\omega \ll \omega_c$.

Figure 32.57 Example 32.8.



❶ **GETTING STARTED** This circuit is the same as the one in Figure 32.52, which I used to determine expressions for V_R (Eq. 32.35) and V_C (Eq. 32.36) in terms of R and C, so I can use those results. From Figure 32.57 I see that the

QUANTITATIVE TOOLS

potential difference v_{out} is equal to the potential difference across the resistor, so $V_{out} = V_R$.

2 DEVISE PLAN In order to determine the value of ω_c at which $V_R = V_C$, I equate the right sides of Eqs. 32.35 and 32.36. The resulting ω factor in my expression then is the cutoff value ω_c. For part *b*, I know that $V_{out} = V_R$. Therefore I can use Eq. 32.35 to determine V_{out} and then determine how V_{out} behaves in the limiting cases where $\omega \gg \omega_c$ and $\omega \ll \omega_c$.

3 EXECUTE PLAN (*a*) Equating the right sides of Eqs. 32.35 and 32.36, I get $R = 1/\omega C$. Solving for ω yields the desired cutoff angular frequency ω_c:

$$\omega_c = \frac{1}{RC}. \checkmark$$

(*b*) To obtain the values of V_{out} for $\omega \gg \omega_c$ and for $\omega \ll \omega_c$, I first rewrite Eq. 32.35 in a form that contains ω_c:

$$V_{out} = V_R = \frac{\mathscr{E}_{max} R}{\sqrt{R^2 + 1/\omega^2 C^2}} = \frac{\mathscr{E}_{max}}{\sqrt{1 + 1/R^2\omega^2 C^2}} = \frac{\mathscr{E}_{max}}{\sqrt{1 + \omega_c^2/\omega^2}}. \quad (1)$$

For $\omega \gg \omega_c$, the second term in the square root is equal to zero and Eq. 1 reduces to $V_{out} = \mathscr{E}_{max}$. ✔

For $\omega \ll \omega_c$, the second term in the square root dominates, so I can ignore the first term. Equation 1 then becomes

$$V_{out} = V_R = \frac{\mathscr{E}_{max}}{\sqrt{1 + \omega_c^2/\omega^2}} \approx \frac{\mathscr{E}_{max}}{\sqrt{\omega_c^2/\omega^2}} = \frac{\mathscr{E}_{max}\omega}{\omega_c} = \mathscr{E}_{max}\omega RC.$$

In the limit that the angular frequency ω approaches zero, V_{out} approaches zero as well. ✔

4 EVALUATE RESULT The name *high-pass filter* makes sense because this circuit allows emfs with an angular frequency higher than the cutoff angular frequency to pass through to the output but attenuates emfs of angular frequency lower than the cutoff angular frequency, preventing them from passing through to the output. It is the capacitor that does the actual passing or blocking. It blocks low-angular-frequency emfs because for these emfs the capacitive reactance, $X_C = 1/\omega C$, is very high. For high-angular-frequency emfs, X_C approaches zero, and so the capacitor passes the emf undiminished.

32.14 Interchange the resistor and the capacitor in Figure 32.57, and then show that the high-pass filter becomes a low-pass filter.

Filters can also be constructed by wiring an inductor and a resistor in series with an AC source. Such a circuit is called an *RL* series circuit and can be analyzed in exactly the manner we used to analyze an *RC* series circuit (as we shall see in Example 32.9).

Finally, let's analyze an *RLC* series circuit: a resistor, a capacitor, and an inductor all in series with an AC source (**Figure 32.58**). As with the *RC* series circuit, the instantaneous current i is the same in all three elements, and the sum of all the potential differences equals the emf of the source:

$$\mathscr{E} = v_R + v_L + v_C. \quad (32.39)$$

The phasor diagram for this circuit is constructed in **Figure 32.59** for the case where $V_L > V_C$. As before, we begin with the phasors for i and v_R,

Figure 32.58 An *RLC* series circuit, consisting of a resistor, an inductor, and a capacitor in series across the terminals of an AC source.

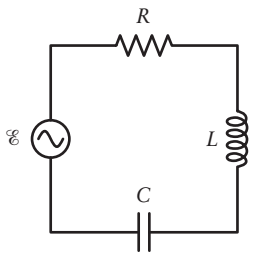
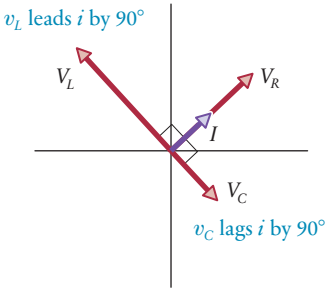
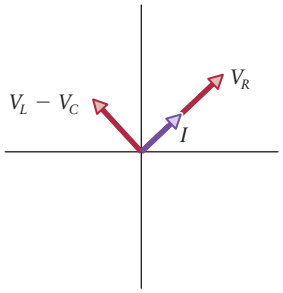


Figure 32.59 Steps involved in constructing a phasor diagram for the *RLC* series circuit in Figure 32.58. The diagram in part *c* indicates the phase of the current relative to the source emf.

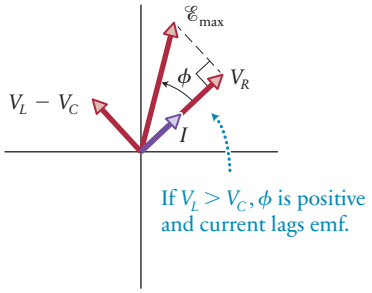
(*a*) Begin with phasors for i and v_R (in phase)



(*b*) Add V_C and V_L



(*c*) Add $V_L - V_C$ and V_R to obtain $\mathscr{E}_{max}$



and then note that v_C lags i by 90° and v_L leads i by 90° (Figure 32.59*a*). As a result, the phasors V_C and V_L can be added directly (Figure 32.59*b*). Finally, the loop rule (Eq. 32.39) requires the phasor for the emf to equal the vector sum of the phasors for the potential differences, as shown in Figure 32.59*c*. Consequently, the amplitudes V_R, V_L, and V_C must satisfy

$$\mathscr{E}_{\max}^2 = V_R^2 + (V_L - V_C)^2. \tag{32.40}$$

Rewriting Eq. 32.40 in terms of I, R (from Eq. 32.6), X_L (from Eq. 32.27), and X_C (from Eq. 32.15) gives

$$\mathscr{E}_{\max}^2 = I^2[R^2 + (X_L - X_C)^2] = I^2[R^2 + (\omega L - 1/\omega C)^2], \tag{32.41}$$

and thus

$$I = \frac{\mathscr{E}_{\max}}{\sqrt{R^2 + (\omega L - 1/\omega C)^2}}. \tag{32.42}$$

The impedance of the RLC series combination (in other words, the constant of proportionality between I and $\mathscr{E}_{\max}$) is therefore

$$Z_{RLC} \equiv \sqrt{R^2 + (\omega L - 1/\omega C)^2} \quad (RLC \text{ series combination}). \tag{32.43}$$

Table 32.1 lists the impedances of various loads.

Figure 32.59*c* shows that the phase relationship between the current and the source emf depends on the relative magnitudes of V_L and V_C. The phase of the current relative to the emf is given by

$$\tan\phi = \frac{V_L - V_C}{V_R} = \frac{X_L - X_C}{R}$$

$$= \frac{\omega L - 1/\omega C}{R} \quad (RLC \text{ series circuit}). \tag{32.44}$$

If $V_L > V_C$, as it is in Figure 32.59, ϕ is positive, meaning that the current lags the source emf. Here the inductor dominates the capacitor, and as a result the series combination of the inductor and capacitor behaves like an inductor. If $V_L < V_C$, ϕ is negative, the inductor-capacitor combination is dominated by the capacitor, and the current leads the source emf, just as in an RC series circuit.

In general, when analyzing AC series circuits, follow the procedure shown in the Procedure box "Analyzing AC series circuits."

Table 32.1 Impedances of various types of loads (all elements in series)

Load	Z
R	R
L	ωL
C	$1/\omega C$
RC	$\sqrt{R^2 + (1/\omega C)^2}$
RLC	$\sqrt{R^2 + (\omega L - 1/\omega C)^2}$

Note that impedances do not simply add the way resistances do. However, the impedance of any simpler load can be found from the impedance of the RLC combination; for example, $Z_{RC} = Z_{RLC}$ without the term containing L.

PROCEDURE: Analyzing AC series circuits

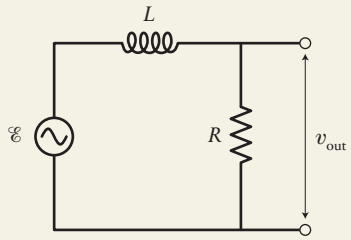
When analyzing AC series circuits, we generally know the properties of the various circuit elements (such as R, L, C, and $\mathscr{E}$) but not the potential differences across them. To determine these, follow this procedure:

1. To develop a feel for the problem and to help you evaluate the answer, construct a phasor diagram for the circuit.
2. Use Eq. 32.43 to determine the impedance of the load. If there is no inductor, then ignore the term containing L; if there is no capacitor, ignore the term containing C; and so on.
3. To determine the amplitude of the current in the circuit, you can now use Eq. 32.42; to determine the phase of the current relative to the emf, use Eq. 32.44.
4. Determine the amplitude of the potential difference across any reactive element using $V = XI$, where X is the reactance of that element. For a resistor, use $V = RI$.

EXAMPLE 32.9 *RL* series circuit

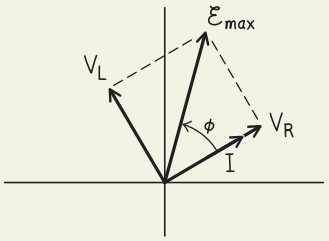
Consider the circuit shown in **Figure 32.60**. (*a*) Determine the cutoff angular frequency ω_c and the phase constant at which $V_R = V_L$. (*b*) Can this circuit be used as a low-pass or high-pass filter?

Figure 32.60 Example 32.9.



❶ **GETTING STARTED** This example is similar to Example 32.8, with the capacitor of that example replaced by an inductor here. As in Example 32.8, I see from the circuit diagram that the potential difference v_{out} is equal to the potential difference across the resistor, so $V_{out} = V_R$. I begin by drawing a phasor diagram for the circuit (**Figure 32.61**). I first draw phasors V_R and I, which I know from Figure 32.45*a* are in phase. I then add V_L, which leads I by 90° (Figure 32.51*a*). I make V_L have the same length as V_R because the problem asks about the circuit when $V_R = V_L$.

Figure 32.61



❷ **DEVISE PLAN** To determine the potential difference amplitudes V_L and V_R across the inductor and the resistor, I follow the procedure given in the Procedure box above. I then set these two amplitudes equal to each other in order to determine ω_c and the phase constant. To determine whether this circuit can be used as a low-pass or high-pass filter, I examine the behavior of V_{out} for $\omega \gg \omega_c$ and for $\omega \ll \omega_c$.

❸ **EXECUTE PLAN** (*a*) Ignoring the term containing C in Eq. 32.43 and substituting the result in Eq. 32.33, I get for the current amplitude

$$I = \frac{\mathscr{E}_{max}}{\sqrt{R^2 + (\omega L)^2}}.$$

I can now use Eq. 32.6 to calculate the amplitude of the potential difference across the resistor,

$$V_R = IR = \frac{\mathscr{E}_{max} R}{\sqrt{R^2 + (\omega L)^2}}, \tag{1}$$

and Eq. 32.27 to calculate the amplitude of the potential difference across the inductor,

$$V_L = IX_L = \frac{\mathscr{E}_{max} \omega L}{\sqrt{R^2 + (\omega L)^2}}, \tag{2}$$

where I have substituted ωL for X_L (Eq. 32.26). Equating the right sides of Eqs. 1 and 2 yields $R = \omega L$. Substituting ω_c for ω in this equation and solving for ω_c give me for the cutoff angular frequency value at which $V_R = V_L$:

$$\omega_c = \frac{R}{L}. \checkmark$$

To determine the phase constant for the condition $V_R = V_L$, I substitute V_R for V_L in Eq. 32.44 and set V_C equal to zero:

$$\tan\phi = \frac{V_L - V_C}{V_R} = \frac{V_R}{V_R} = 1,$$

so the phase constant is 45°. ✔

(*b*) Just as I did in Example 32.8, to obtain the limiting values of V_{out}, I first rewrite Eq. 1 in a form that contains ω_c:

$$V_{out} = V_R = \frac{\mathscr{E}_{max} R}{\sqrt{R^2 + (\omega L)^2}} = \frac{\mathscr{E}_{max}}{\sqrt{1 + (\omega L)^2/R^2}}$$

$$= \frac{\mathscr{E}_{max}}{\sqrt{1 + \omega^2/\omega_c^2}}. \tag{3}$$

(Continued)

For $\omega \ll \omega_c$, the second term in the square root is equal to zero and Eq. 3 reduces to $V_{out} = \mathscr{E}_{max}$. For $\omega \gg \omega_c$ the second term in the square root dominates and we can ignore the first term. Equation 3 then becomes

$$V_{out} = V_R = \frac{\mathscr{E}_{max}}{\sqrt{1 + \omega^2/\omega_c^2}} \approx \frac{\mathscr{E}_{max}}{\sqrt{\omega^2/\omega_c^2}} = \frac{\mathscr{E}_{max}\omega_c}{\omega} = \frac{\mathscr{E}_{max}R}{\omega L}.$$

In the limit that the angular frequency ω becomes very large, V_{out} approaches zero. The circuit thus blocks high-frequency emfs and allows low-frequency ones to pass through to the output. Therefore it can be used as a low-pass filter. ✔

4 EVALUATE RESULT From my phasor diagram I see that the triangle that has V_R and V_L as two of its sides is an equilateral right-angle triangle, and so the phase constant ϕ must be 45°, as I obtained.

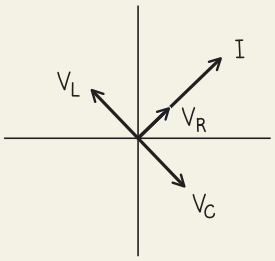
For part *b*, an emf is generated in the inductor whenever the current in it changes. This emf is proportional to the rate of change of the current in the inductor and it opposes the change in current (Eq. 29.19, $\mathscr{E}_{ind} = -L(di/dt)$). For a low-angular-frequency emf, di/dt is small, and the signal passes through the inductor essentially undiminished. For a high-angular-frequency signal, the inductive reactance $X_L = \omega L$ is high, and so the inductor essentially blocks the signal. It therefore makes sense that the arrangement in Figure 32.60 can serve as a low-pass filter.

EXAMPLE 32.10 *RLC* series circuit

Consider an *RLC* circuit, such as the one shown in Figure 32.58. The source emf has amplitude 160 V and frequency 60 Hz. The resistance is $R = 50\ \Omega$ and the inductance is $L = 0.26$ H. If the amplitudes of the potential difference across the capacitor and the inductor are equal, what is the current in the circuit?

1 GETTING STARTED I begin by drawing a phasor diagram for the circuit (**Figure 32.62**). I first draw phasors V_R and I, which are in phase, arbitrarily choosing the direction in which I draw them. I then add phasors V_C, which lags I by 90°, and V_L, which leads I by 90°.

Figure 32.62



2 DEVISE PLAN The current in the circuit depends on the impedance, which is given by Eq. 32.43. This equation contains *C*, however, and I am given no information about this variable. I am given, however, that $V_C = V_L$, and because both V_C and V_L are proportional to the current, I should be able to determine the current without knowing the capacitance in the circuit.

3 EXECUTE PLAN I know from Eq. 32.24 that $V_L = I\omega L$, and I also know that $V_C = I/\omega C$ (Eq. 32.13). Because $V_L = V_C$ in this problem, I can equate the terms on the right in these two equations to obtain

$$\omega L = \frac{1}{\omega C}. \qquad (1)$$

Substituting ωL for $1/\omega C$ in Eq. 32.43 yields $Z_{RLC} = R$. Now I can use Eq. 32.33 to determine the current in the circuit:

$$I = \frac{\mathscr{E}_{max}}{Z_{RLC}} = \frac{\mathscr{E}_{max}}{R} = \frac{160\ \text{V}}{50\ \Omega} = 3.2\ \text{A}. \checkmark$$

4 EVALUATE RESULT When the amplitudes of the potential differences across the inductor and the capacitor are equal, the lengths of the phasors V_C and V_L are equal. Because the phasors point in opposite directions (Figure 32.62), they add to zero, and so the impedance in the circuit is due to the resistor only. This means the current is essentially given by Ohm's law, $I = V/R$, or, in the version I obtained here, $I = \mathscr{E}_{max}/R$.



32.15 (*a*) Calculate the maximum potential difference across each of the three circuit elements in Example 32.10. (*b*) Is the sum of the amplitudes V_R, V_L, and V_C equal to the amplitude of the source emf? Why or why not?

32.7 Resonance

Consider again the *RLC* series circuit of Figure 32.58. Suppose that the amplitude $\mathscr{E}_{max}$ of the source emf is held constant, but we vary its angular frequency. What happens to the amplitude of the current I and to the phase constant ϕ? Combining Eqs. 32.42 and Eq. 32.43, we can say

$$I = \frac{\mathscr{E}_{max}}{Z} = \frac{\mathscr{E}_{max}}{\sqrt{R^2 + (\omega L - 1/\omega C)^2}}. \qquad (32.45)$$

The current is at its maximum when the term in parentheses in the denominator is zero. When this term is zero, $V_L = V_C$ and the two potential

QUANTITATIVE TOOLS

differences are 180° out of phase. Therefore the effects of the inductor and the capacitor cancel each other and the circuit behaves as if only the resistor is present. The term in parentheses in Eq. 32.45 is zero when

$$\omega L = \frac{1}{\omega C}. \tag{32.46}$$

The angular frequency for which Eq. 32.46 is satisfied is called the **resonant angular frequency** ω_0 of the circuit:

$$\omega_0 = \frac{1}{\sqrt{LC}}. \tag{32.47}$$

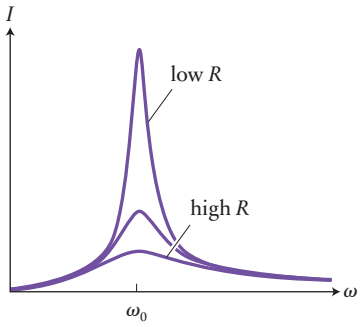
The current amplitude and phase as a function of angular frequency are plotted in **Figure 32.63** for three values of R (with fixed values of $\mathscr{E}_{max}$, L, and C). Increasing or decreasing the angular frequency from ω_0 decreases the current amplitude. Changing R changes the maximum current that can be obtained and also changes how rapidly the current drops as the angular frequency increases or decreases from resonance.

Whenever an oscillating physical quantity has a peaked angular frequency dependence, the dependence is referred to as a *resonance curve*. The sharpness of the peak reflects the efficiency with which the source delivers energy to the system at or near resonance and depends on the amount of dissipation present in the system. A very tall, sharp peak corresponds to a system with low dissipation. In such a system, the source can pump an enormous amount of energy into the system at resonance. A short, broad peak corresponds to a system with high dissipation. Here, less energy goes into the system even at resonance, but that energy can be transferred in at angular frequencies farther from resonance. For the RLC series circuit, energy is dissipated via the resistor; high R values produce less current at ω_0 and a broader resonance curve, as Figure 32.63*a* shows.

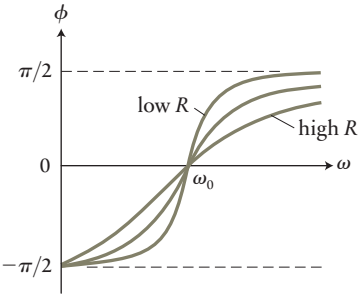
Another system that exhibits resonance is a damped mechanical oscillator (see Section 15.8) driven by an external source. The damping in a mechanical oscillator is analogous to the resistance in the RLC series circuit.

Figure 32.63 Current and phase changes in the RLC circuit of Figure 32.58.

(*a*) Frequency dependence of the current at low, medium, and high R



(*b*) Frequency dependence of current phase relative to source emf as function of angular frequency at low, medium, and high R



32.16 How does the resonance curve in Figure 32.63 change if the value of C or L is changed?

In the RLC series circuit, the phase difference between the current and the driving emf also depends on the angular frequency of the AC source. The current can either lag or lead the emf (or be in phase with it), depending on the angular frequency. We found previously that the phase of the current relative to the source emf for an RLC series circuit is given by Eq. 32.44:

$$\tan\phi = \frac{\omega L - 1/\omega C}{R}. \tag{32.48}$$

Consider the limiting values of this expression for the relative phase by looking at the curves in Figure 32.63*b*. At resonance ($\omega = \omega_0$), $\phi = 0$ and the current and the source emf are in phase. When $\omega = 0$, $\tan\phi = -\infty$ and $\phi = -\pi/2$. When $\omega = \infty$, $\tan\phi = \infty$ and $\phi = +\pi/2$. Below resonance, $\phi < 0$, the capacitor provides the dominant contribution to the impedance, and the current leads the source emf. Above resonance, $\phi > 0$ and the inductor dominates, and the current lags the source emf.

32.17 In an RLC series circuit, you measure $V_R = 4.9$ V, $V_L = 6.7$ V, and $V_C = 2.5$ V. Is the angular frequency of the AC source above or below resonance?

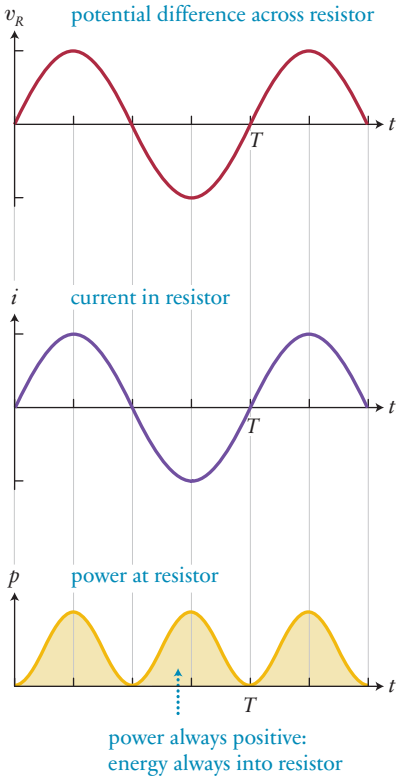


Figure 32.64 For an AC circuit consisting of a resistor connected across an AC source, time dependence of potential difference across the resistor, current in the resistor, and power dissipated at the resistor.

32.8 Power in AC circuits

At the beginning of this chapter, we saw that in alternating-current circuits, the energy stored in capacitors and inductors can oscillate. Consequently, for part of each cycle, these elements put energy back into the source rather than taking up energy from the source. Thus, unlike what we see in DC circuits, the source in an AC circuit does not simply deliver energy steadily to the circuit. Let's take a closer look at how to determine the rate at which an AC source delivers energy to a load.

In general, the rate at which the source delivers energy to its load—in other words, the power of the source—is the time-dependent version of the result we found for DC circuits (Eq. 31.42):

$$p = iv_{\text{load}}. \tag{32.49}$$

Because the current and the emf oscillate, this power varies with time and in principle can be either positive or negative. Let's first consider a load that consists of just one resistor. Ohm's law tells us that the instantaneous energy delivered to the resistor is

$$p = iv_R = i^2R = I^2R \sin^2 \omega t. \tag{32.50}$$

The time dependence of the potential difference, current, and power are shown in **Figure 32.64**. Because the current and potential difference are in phase, the power is always positive, and so the source always delivers energy to the resistor. This makes sense because the resistor dissipates energy regardless of the current direction. Consequently, the rate at which energy is dissipated in the resistor (the *power at the resistor*) is always positive and oscillates at *twice* the angular frequency of the emf.

For most applications, we are interested in the time average of the power at the resistor. Using the trigonometric identity $\sin^2\alpha = \frac{1}{2}(1 - \cos 2\alpha)$, we can rewrite Eq. 32.50 as

$$p = I^2R[\tfrac{1}{2}(1 - \cos 2\omega t)] = \tfrac{1}{2} I^2R - \tfrac{1}{2} I^2R \cos 2\omega t. \tag{32.51}$$

The first term on the right is constant in time. The second term on the right averages to zero over a full cycle because the area under the positive half of the cosine is equal to the area under the negative half. As a result, for time intervals much longer than the period of oscillation, the time average of the power at the resistor is

$$P_{\text{av}} = \tfrac{1}{2} I^2R. \tag{32.52}$$

For a sinusoidally varying current, the *root-mean-square* or *rms* value of the current is (Eqs. 19.21 and 30.39)

$$I_{\text{rms}} \equiv \sqrt{(i^2)_{\text{av}}} = \sqrt{\tfrac{1}{2}I^2} = \frac{I}{\sqrt{2}}. \tag{32.53}$$

and so

$$P_{\text{av}} = I_{\text{rms}}^2 R. \tag{32.54}$$

The advantage of writing the average power in terms of the rms current is that Eq. 32.54 is completely analogous to the expression for the rate of energy dissipation by a resistor connected to a DC source (Eq. 31.43). Similarly, we can introduce rms values of potential difference and source emf:

$$V_{\text{rms}} = \frac{V_R}{\sqrt{2}} \quad \text{and} \quad \mathscr{E}_{\text{rms}} = \frac{\mathscr{E}_{\text{max}}}{\sqrt{2}}. \tag{32.55}$$

QUANTITATIVE TOOLS

The rms value is a useful way to measure the average value of an oscillating current or emf because the strict time average of these quantities is zero. *Voltmeters* and *ammeters* typically measure the rms value of alternating potential differences and currents, respectively. Thus, for example, the wall potential difference in household electrical wiring in the United States is referred to as 120 V even though the amplitude is 170 V; the 120-V rating is the rms value.

Next let's look at a circuit made up of an AC source and a capacitor. How much power is delivered to the capacitor by the source? Now the current and the potential difference are out of phase (see Figure 32.48). Substituting from Eq. 32.8, $v_C = V_C \sin \omega t$, and Eq. 32.11, $i = \omega C V_C \cos \omega t$, into Eq. 32.49, we obtain

$$p = iv_C = (\omega C V_C \cos \omega t)(V_C \sin \omega t). \tag{32.56}$$

Using the trigonometric identity $\sin(2\alpha) = 2 \sin \alpha \cos \alpha$, we obtain

$$p = \tfrac{1}{2}\omega C V_C^2 \sin 2\omega t. \tag{32.57}$$

As in the resistor-only circuit, the power oscillates at twice the angular frequency of the source, but now the power is sometimes positive and sometimes negative, as shown in **Figure 32.65**. When v and i have the same sign, energy is transferred to the capacitor; when v and i have opposite signs, energy residing in the capacitor is transferred back to the source. The average power is zero, as it must be, because no energy is dissipated in a capacitor. The same is true for an inductor, except that in an inductor energy is stored as magnetic energy rather than electric energy; you can show this mathematically by converting Eqs. 32.56 and 32.57 to their V_L counterparts.

Finally, let's examine how power is delivered to the load in an *RLC* circuit. Although we could work out the power at each element, it's easier to consider the power for the entire load consisting of the *RLC* combination. The potential difference across the load is equal to the applied emf. Using Eq. 32.1, $\mathscr{E} = \mathscr{E}_{\max} \sin \omega t$, and Eq. 32.16, $i = I \sin(\omega t - \phi)$, we can write the instantaneous power as

$$p = vi = \mathscr{E}_{\max} I \sin \omega t \sin(\omega t - \phi). \tag{32.58}$$

Using the trigonometric identity $\sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta$ to separate the ϕ and ω dependence and substituting $\sin \alpha \cos \alpha = \frac{1}{2} \sin 2\omega t$, we rewrite Eq. 32.58 in the form

$$\begin{aligned} p &= \mathscr{E}_{\max} I (\cos \phi \sin^2 \omega t - \sin \phi \sin \omega t \cos \omega t) \\ &= \mathscr{E}_{\max} I (\cos \phi \sin^2 \omega t - \sin \phi \tfrac{1}{2} \sin 2\omega t). \end{aligned} \tag{32.59}$$

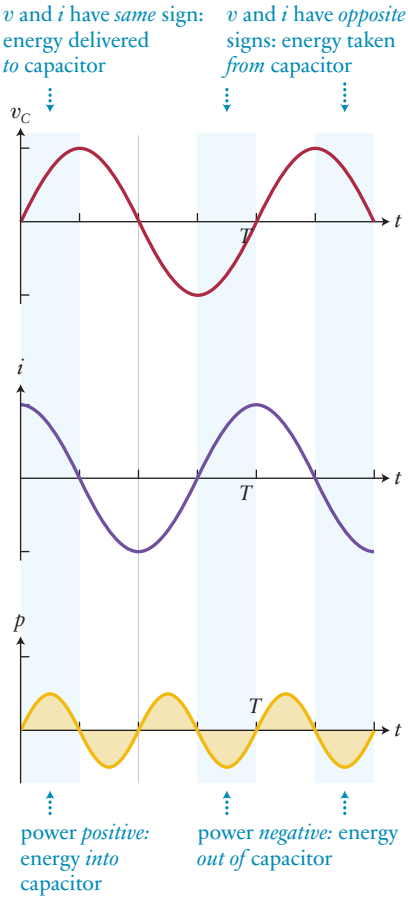
The time average of $\sin^2 \omega t$ is 1/2, and the second term inside the parentheses averages to zero, leaving us with

$$P_{\text{av}} = \tfrac{1}{2} \mathscr{E}_{\max} I \cos \phi. \tag{32.60}$$

Writing this result using rms values (Eqs. 32.54 and 32.55) gives

$$P_{\text{av}} = \tfrac{1}{2}(\sqrt{2}\, \mathscr{E}_{\text{rms}})(\sqrt{2}\, I_{\text{rms}}) \cos \phi = \mathscr{E}_{\text{rms}} I_{\text{rms}} \cos \phi. \tag{32.61}$$

Figure 32.65 For an AC circuit consisting of a capacitor connected across an AC source, time dependence of potential difference across the capacitor, current in the capacitor, and power into or out of the capacitor.



We can rewrite this in a more physically insightful way if we note that $\mathscr{E}_{\text{rms}} = I_{\text{rms}} Z$ (Eq. 32.33) and note from Figure 32.59*c* that

$$\cos \phi = \frac{V_R}{\mathscr{E}_{\text{max}}} = \frac{RI}{ZI} = \frac{R}{Z}. \tag{32.62}$$

With these substitutions, Eq. 32.61 becomes

$$P_{\text{av}} = I_{\text{rms}} Z \, I_{\text{rms}} \frac{R}{Z} = I_{\text{rms}}^2 \text{R}. \tag{32.63}$$

This result tells us that all of the energy delivered to the circuit is dissipated as thermal energy in the resistor—as it must be, because neither the capacitor nor the inductor dissipates energy. This energy is dissipated at the same average rate as in a circuit made up of a single resistor connected to an AC source (Eq. 32.54).

The factor $\cos \phi$ that appears in Eqs. 32.60–32.62 is called the **power factor** which is a measure of the efficiency with which the source delivers energy to the load. At resonance, when the current and the emf are in phase ($\phi = 0$), the current and the power factor are greatest, and the maximum power possible is delivered to the load. At angular frequencies away from resonance, less power is delivered to the load.



32.18 Calculate the rate P_{av} at which energy is dissipated in the *RLC* series circuit of Example 32.10.

Chapter Summary

AC circuits (Sections 32.1, 32.2, 32.5)

Concepts **Alternating current** (AC) is current that periodically changes direction.

In an AC circuit that contains a capacitor, the current through the capacitor leads the potential difference by 90° (one-fourth of an oscillation).

In an AC circuit that contains an inductor, the current through the inductor lags the potential difference by 90°.

In any circuit that contains a capacitor and/or an inductor, **reactance** is the proportionality constant between the potential-difference amplitude and the current amplitude.

The **phase constant** ϕ for an AC circuit, which is the phase difference between the source emf and the current, is negative when the current leads the source emf and positive when the current lags the source emf.

Quantitative Tools The instantaneous emf supplied by an AC source is

$$\mathscr{E} = \mathscr{E}_{\max} \sin \omega t. \tag{32.1}$$

The instantaneous current in an AC circuit is

$$i = I \sin(\omega t - \phi), \tag{32.16}$$

where I is the amplitude (maximum value) of the current and ϕ is the **phase constant**.

In an AC circuit that contains a capacitor, the *capacitive reactance* is

$$X_C \equiv \frac{1}{\omega C}, \tag{32.14}$$

and the amplitude of the potential difference across the capacitor is

$$V_C = IX_C. \tag{32.15}$$

In an AC circuit that contains an inductor, the *inductive reactance* is

$$X_L \equiv \omega L, \tag{32.26}$$

and the amplitude of the potential difference across the inductor is

$$V_L = IX_L. \tag{32.27}$$

RC and *RLC* series circuits (Section 32.6)

Concepts The **impedance** of a load connected to an AC source is a proportionality constant between the amplitudes of the potential difference and the current through the load.

Quantitative Tools In an ***RC* series circuit,** the sum of all the instantaneous potential differences equals the emf of the source:

$$\mathscr{E} = v_R + v_C. \tag{32.28}$$

The current amplitude is

$$I = \frac{\mathscr{E}_{\max}}{Z}, \tag{32.33}$$

where Z is the **impedance** of the load. For an RC series circuit, the impedance is

$$Z_{RC} \equiv \sqrt{R^2 + 1/\omega^2 C^2}. \tag{32.34}$$

The amplitudes of the potential differences are

$$V_R = IR = \frac{\mathscr{E}_{\max} R}{\sqrt{R^2 + 1/\omega^2 C^2}} \tag{32.35}$$

$$V_C = IX_C = \frac{\mathscr{E}_{\max}/\omega C}{\sqrt{R^2 + 1/\omega^2 C^2}}, \tag{32.36}$$

and the **phase constant** for the circuit is

$$\phi = \tan^{-1}\left(-\frac{1}{\omega RC}\right). \tag{32.38}$$

In an ***RLC* series circuit,** the sum of all the instantaneous potential differences equals the emf of the source:

$$\mathscr{E} = v_R + v_L + v_C. \tag{32.39}$$

The **impedance** of the *RLC* series circuit is

$$Z_{RLC} \equiv \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}, \tag{32.43}$$

and so the current amplitude is

$$I = \frac{\mathscr{E}_{\max}}{Z_{RLC}} = \frac{\mathscr{E}_{\max}}{\sqrt{R^2 + (\omega L - 1/\omega C)^2}}. \tag{32.42}$$

The **phase constant** for the circuit is given by

$$\tan\phi = \frac{\omega L - 1/\omega C}{R}. \tag{32.44}$$

Resonance and power in AC circuits (Sections 32.7, 32.8)

Concepts At the **resonant angular frequency,** the current in a series *RLC* circuit is a maximum.

The **power factor** is a measure of the efficiency with which the source in an AC circuit delivers energy to the load.

Quantitative Tools The **resonant angular frequency** is

$$\omega_0 = \frac{1}{\sqrt{LC}}. \tag{32.47}$$

For a sinusoidally varying current, the *root-mean-square* current is

$$I_{\text{rms}} = \frac{I}{\sqrt{2}}, \tag{32.53}$$

and the root-mean-square potential difference is

$$V_{\text{rms}} = \frac{V}{\sqrt{2}} \qquad \mathscr{E}_{\text{rms}} = \frac{\mathscr{E}_{\max}}{\sqrt{2}}. \tag{32.55}$$

The average power is

$$P_{\text{av}} = \mathscr{E}_{\text{rms}} I_{\text{rms}} \cos\phi, \tag{32.61}$$

where the cosine term is the **power factor** for the circuit:

$$\cos\phi = \frac{R}{Z}. \tag{32.62}$$

Semiconductor devices (Sections 32.3, 32.4)

Concepts A **semiconductor** is a material that contains a limited number of charge carriers that can move freely and that has an electrical conductivity intermediate between that of electrical conductors and that of electrical insulators.

An *intrinsic* semiconductor is made of atoms of only one element. An *extrinsic* (*doped*) semiconductor contains, interspersed among the atoms of its main element, trace amounts of atoms of other elements. These *dopant* atoms alter the number of electrons that can move freely and so change the electronic properties of the semiconductor.

A **hole** is an incomplete bond in a semiconductor that behaves like a freely moving positive charge carrier.

In a *p-type* semiconductor, holes are the free charge carriers. In an *n-type* semiconductor, electrons are the free charge carriers.

A **diode** is made by bringing a piece of *p*-type semiconductor into contact with a piece of *n*-type semiconductor. It behaves like a one-way valve for current.

The **depletion zone** in a diode is a thin insulating region at the junction between the *p*-type and *n*-type semiconductors where the positive and negative charge carriers have recombined and become immobile.

A **transistor** consists of a thin layer of one type of extrinsic semiconductor sandwiched between two layers of the opposite type of extrinsic semiconductor (such as *npn*). It behaves either like an on/off switch for current or like a current amplifier.

Questions and Problems

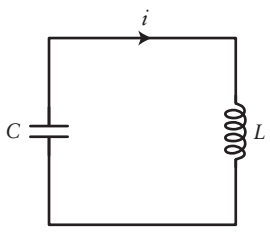
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

32.1 Alternating currents

1. A fully charged parallel-plate capacitor is connected to an inductor with n turns of wire (Figure P32.1) What is the rms value of the magnetic field through the inductor, which has an rms current of I_{rms}? •

Figure P32.1

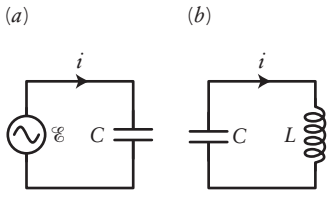


2. (*a*) If each vertical gray line in Figure 32.4*b* represents 0.020 s, what is the frequency of the oscillation? (*b*) What is the average potential difference across the capacitor in one cycle? •

3. For the circuit shown in Figure 32.4*a*, let the period T of the oscillation be 1.0 s and the emf be zero at $t = 0$. Identify the instant(s) at which the energy stored in the inductor reaches a maximum. •

4. (*a*) For the circuit shown in Figure P32.4*a*, draw (qualitative) curves showing the current as a function of time and the potential difference across the parallel-plate capacitor as a function of time. Assume that the capacitor is fully

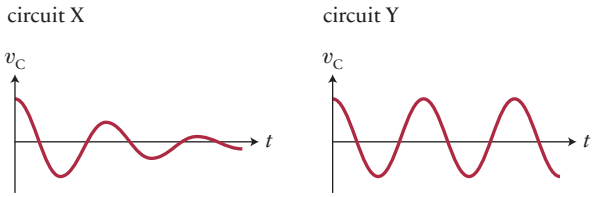
Figure P32.4



charged at $t = 0$ and that the potential is positive when the upper plate is positively charged, and use the current reference direction shown. (*b*) Repeat for the circuit in Figure P32.4*b*. ●●

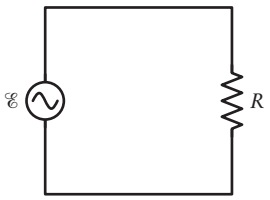
5. Two circuits X and Y each contain a parallel-plate capacitor and may contain other elements as well. The capacitors are identical, and the potential difference across each is shown in Figure P32.5. Which circuit—X, Y, or both—could contain a resistor, and under what circumstances could it/they contain a resistor? ●●

Figure P32.5



6. In an *LC* circuit like the one shown in Figure P32.1, which of these quantities simultaneously reach their maximum values: $|v_C|$, $|q_C|$, $|i|$, $|B|$, $|U^E|$, $|U^B|$? ●●
7. A parallel-plate capacitor is fully charged and connected in series to an inductor only. The capacitor initially stores charge q and electric potential energy U^E. Once the circuit is closed, charge carriers flow from the capacitor, producing a current in the circuit. When the charge on the capacitor has been reduced to $q/2$, what is the magnetic potential energy U^B stored in the inductor? Report your answer in terms of U^E. ●●
8. If the circuit in Figure P32.8 operates at 50 Hz with $\mathscr{E}_{max} = 120$ V and $R = 15.0\ \Omega$, how much energy is dissipated in the resistor in 1.5 s? ●●

Figure P32.8

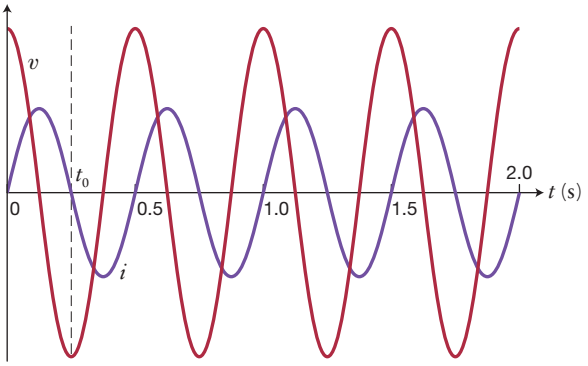


9. A circuit consists of a parallel-plate capacitor, an inductor, and a switch, which is initially open. The capacitor plates have an area of $1.00 \times 10^{-4}\ m^2$ and are held 0.100 mm apart by a thin layer of barium titanate. The capacitor is charged to a potential difference of 10 V, and then the switch is closed. How much charge remains on the capacitor at the instant the energy stored in the inductor has increased to 85% of its maximum? (See Table 26.1.) ●●●
10. In an *LC* circuit, there is always some energy dissipation as electrical energy is converted to thermal energy. (*a*) What property of the circuit is responsible for electrical energy being converted to thermal energy? (*b*) Describe the shape of a graph showing the charge on one capacitor plate as a function of time. (*c*) What sort of function would describe the rate at which energy is converted? [Hint: See the discussion of damped mechanical oscillators in Chapter 15.] ●●●

32.2 AC circuits

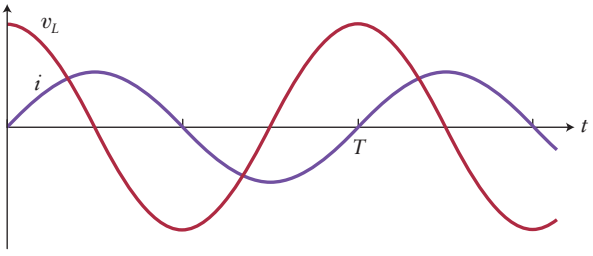
11. Draw a phasor diagram at instant t_0 for the circuit element whose instantaneous current and instantaneous potential difference are shown in Figure P32.11. ●

Figure P32.11



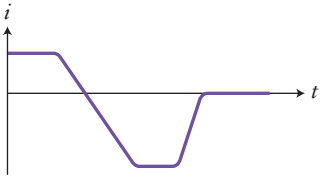
12. Draw a phasor diagram for the current and potential difference represented in Figure P32.12 at the instant $t = T$. Be sure to note which way it is rotating. ●

Figure P32.12



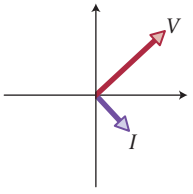
13. In a circuit consisting of an inductor connected to a nonsinusoidal AC source, the current varies as a function of time as shown in Figure P32.13. Sketch the potential difference across the inductor as a function of time. ●●

Figure P32.13



14. The phasor diagram in Figure P32.14 shows the potential difference across a circuit element and the current through the element at $t = 0$. (*a*) Is the element a resistor,

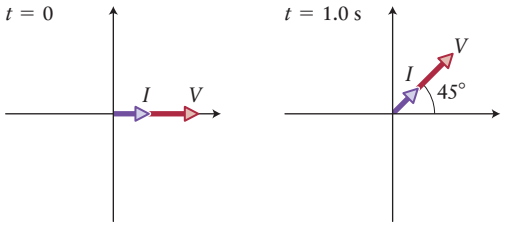
Figure P32.14



capacitor, or inductor? (*b*) On a single graph, sketch a curve showing v as a function of time and a curve showing i as a function of time. ●●

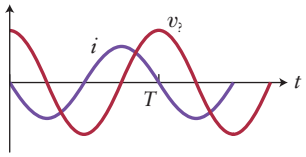
15. Figure P32.15 shows, for an AC circuit, the current phasor and a potential difference phasor at $t = 0$ and at $t = 1.0$ s. In each case, assume phasor magnitudes $I = 1.0$ A and $V = 9.0$ V. Give as many details about this circuit as you can. ●●

Figure P32.15



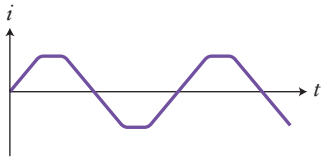
16. Figure P32.16 shows, for a circuit consisting of one element and an AC source, the current through the element as a function of time and the potential difference across the element. Is the element an inductor, a capacitor, or a resistor? ●●

Figure P32.16



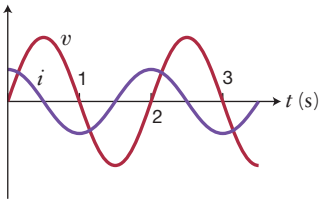
17. In a circuit consisting of a nonsinusoidal AC source plus one additional element, the instantaneous current varies as shown in Figure P32.17. Sketch the potential difference across the additional element as a function of time if that element is (*a*) a capacitor and (*b*) an inductor. ●●●

Figure P32.17



18. A friend has built an AC circuit for you to analyze. You know the circuit contains a resistor, but it also contains another element. Because you know the circuit components available to your friend, you know that the unknown element must be a 0.32-F capacitor, a 0.32-mF capacitor, a 6.28-mH inductor, or a 62.8-mH inductor. Figure P32.18 shows curves for the current through the unknown element and the potential difference across it. The maximum current is 1.00 mA, and the maximum potential difference is 1.00 V. Which of the four available elements did your friend use, and how do you know? ●●●

Figure P32.18



32.3 Semiconductors

19. A sample contains 3.0×10^{22} Si atoms and nothing else. (*a*) How many valence electrons does the sample contain? (*b*) How many of the valence electrons in the sample can conduct electricity? ●

20. Silicon is doped with the following elements. In each case, is the resultant extrinsic semiconductor *n*-type or *p*-type: indium, gallium, phosphorus, boron, arsenic, aluminum, antimony? ●

21. Conduction in a semiconductor can be greatly enhanced by the addition of very small quantities of dopant atoms, even in the parts-per-million range. Explain how the addition of such a tiny percentage of dopant atoms can increase the conduction of a semiconductor by a significant factor. ●

22. Unlike electrons, the holes in a semiconductor do not leave the semiconductor in which they are formed. Yet, holes can be thought of as positive charge carriers. When a hole has traveled to the edge of the semiconductor and reaches an electrical lead, where does it go? ●●

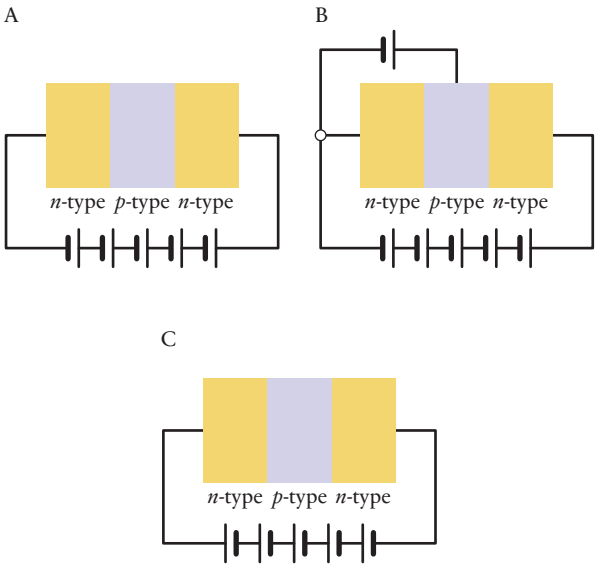
23. You have an intrinsic semiconductor, a *p*-type semiconductor, and an *n*-type semiconductor. The doped samples each contain the same number density of dopant atoms. Which of the three semiconductors do you expect to have (*a*) the greatest ability to carry current when the three are connected to identical batteries, (*b*) the greatest number of electrons, and (*c*) the greatest number of holes entering the battery as they leave the semiconductor? ●●

24. You have ordered a 10.00 mm × 10.00 mm × 2.000 mm sheet of pure silicon from a laboratory supply company, but the sheet the company sends has a mass of 465.9480 mg. You know that the company also produces silicon doped with boron, and you wonder if there may have been some contamination. The mass densities are $\rho_{Si} = 2329.6\ \text{kg/m}^3$ and $\rho_B = 2340\ \text{kg/m}^3$, and the atomic weights are B = 10.81 atomic units and Si = 28.085 atomic units. Determine as much information as possible about the contamination of your sheet. If you need pure silicon for its conduction properties, is this sheet likely to meet your needs? ●●●

32.4 Diodes, transistors, and logic gates

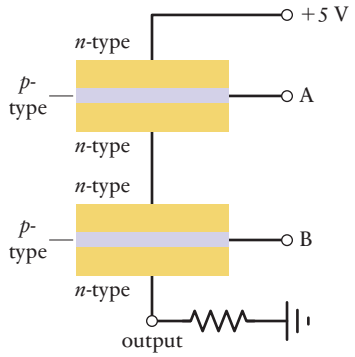
25. Which circuit in Figure P32.25 produces the greatest current in the emitter? •

Figure P32.25



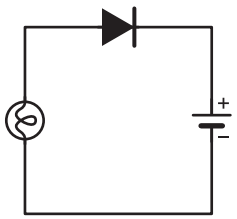
26. The two transistors in Figure P32.26 are connected to each other by wires and also connected to input wires at terminals A and B. The output of this combination depends on the inputs at A and B. In terms of this circuit's reaction to input at A and B, which type of logical operation is performed? •

Figure P32.26



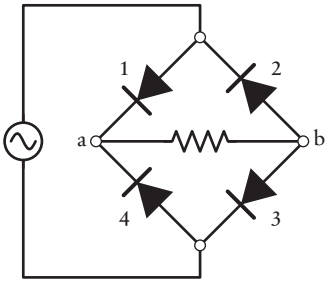
27. Is the light bulb in Figure P32.27 lit? •

Figure P32.27



28. The rectifier shown in Figure P32.28 converts the alternating emf from an AC source to a positive-only potential difference from a to b. Sketch graphs showing (*a*) the source emf as a function of time and (*b*) the potential difference from a to b as a function of time. (*c*) If diode 1 is blown out and acts as an open circuit, what is the output potential difference from a to b? (*d*) What happens to the potential difference from a to b if diode 1 operates normally but diode 2 is blown out? ••

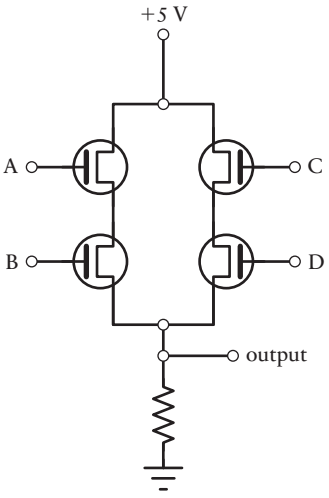
Figure P32.28



29. The circuit diagram for a rectifier is sometimes drawn with a resistor in the center. Would the rectifier function if this resistor were replaced by a capacitor? ••

30. The logic gate in Figure P32.30 contains four field-effect transistors. Each of the gates A–D of the transistors may be positively biased (Y in the table) or not (N in the table). Depending on the various bias combinations, charge carriers may or may not flow through the circuit. Fill in the output column of the table. ••

Figure P32.30



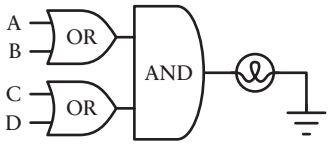
A	B	C	D	output
Y	Y	Y	Y	
Y	N	Y	Y	
Y	Y	Y	N	
N	N	Y	Y	
N	Y	Y	N	
N	N	N	Y	

31. You have two regions, A and B, of differently doped silicon, with the regions joined together to make a continuous silicon crystal. When the positive terminal of a battery is connected to region A and the negative terminal is connected to region B, no charge carriers flow. When the terminal connections are switched, charge carriers flow. (*a*) When there is a current, in which direction—A to B or B to A—do the positive charge carriers move? (*b*) In which direction do the electrons move? (*c*) Which region is *p*-type and which is *n*-type? (*d*) Which region may contain boron, and which may contain phosphorus? ••

32. Suppose a transistor consists of a very narrow *p*-type material sandwiched between two very wide regions of *n*-type material. (*a*) Is the charge on the *p*-type region positive or negative, and why? (*b*) Is the charge on the *n*-type region positive or negative, and why? ••

33. In Figure P32.33, what combinations of positive bias (input signals) A, B, C, D allow the light bulb to light up? ••

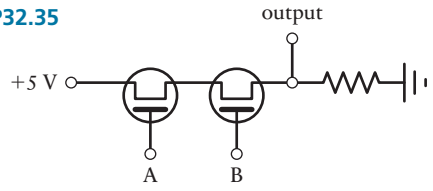
Figure P32.33



34. If in Figure P32.33 the two gates on the left were both AND and the gate on the right was OR, what combinations of positive bias (input signals) A, B, C, D would allow the light bulb to light up? ••

35. You are using the logic gate of Figure P32.35 to test some transistors and emf sources. The output produced by the logic gate depends on the potential relative to ground at inputs A and B. The only emf sources available to you are AC sources, but using one of them to bias either input A or input B results in an oscillating input potential. Consider how using an AC source affects the potential relative to ground at either A or B. Describe how using an AC source differs from using a DC source, both in general and in the specific function of this circuit. Could any modifications or additions be made to the circuit so that the way it functions with an AC source more closely resembles the way it would function with a DC source? Draw a circuit diagram of any useful modifications. •••

Figure P32.35



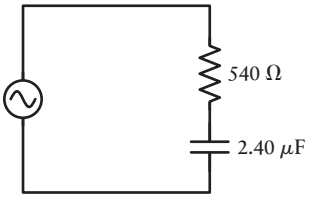
32.5 Reactance

36. In a circuit consisting of a 2 kΩ resistor connected in series with a 50 mH inductor, the amplitude of the emf is $\mathscr{E}_{max} = 20$ V, and a 20-kHz signal is applied. Find impedance, current, phase angle, and potential differences across the resistor and inductor. ••

37. The reactance of a 10-μF capacitor is $X_C = 200\ \Omega$. What is the angular frequency of the current through the capacitor? •

38. For the circuit shown in Figure P32.38, what must the frequency f of the AC source be in order for the amplitude V_C of the potential difference across the capacitor to equal the amplitude V_R of the potential difference across the resistor? •

Figure P32.38

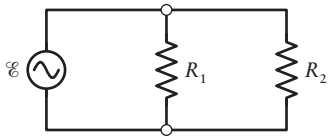


39. A capacitor for which $C = 10$ mF is connected across an AC source. The source emf is given by $\mathscr{E} = \mathscr{E}_{max} \sin \omega t$, with $\mathscr{E}_{max} = 5.0$ V and $\omega = 20\ s^{-1}$. (*a*) What is the amplitude of the current through the capacitor? (*b*) How does the current amplitude change as you increase the source angular frequency? ••

40. An inductor for which $L = 10$ mH is connected to an AC source. The source emf is given by $\mathscr{E} = \mathscr{E}_{max} \sin \omega t$, with $\mathscr{E}_{max} = 5.0$ V and $\omega = 200\ s^{-1}$. (*a*) What is the amplitude of the current through the inductor? (*b*) How does the current amplitude change as you lower the source angular frequency? As you lower the inductance? ••

41. Figure P32.41 shows resistors $R_1 = 10.0\ \Omega$ and $R_2 = 20.0\ \Omega$ connected in parallel to an AC source. (*a*) If the maximum current in the circuit is 12 A, what is the amplitude $\mathscr{E}_{max}$ of the AC source? (*b*) If the source emf peaks at $t = 0$, and the frequency is 60.0 Hz, what is the potential difference across each resistor at $t = 12.5$ ms? ••

Figure P32.41



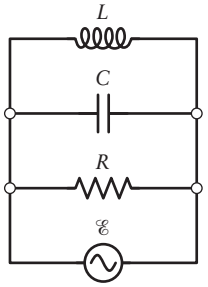
42. You must use a resistor, a capacitor, or an inductor to fill a gap in a circuit. You want to have as much current as possible through this circuit element to ensure that it can be easily detected. Which element should you use if the current through it is (*a*) very high frequency (VHF), (*b*) very low frequency (VLF), (*c*) DC, (*d*) time-varying so that it is sometimes VHF and other times VLF? ••

43. A device is connected to a 60.0-Hz AC source for which the emf is given by $\mathscr{E} = \mathscr{E}_{max} \sin(\omega t)$. If after 5.00 ms the current is 30.9% of its maximum value, is the current leading or lagging the source emf? ••

44. A resistor for which $R = 60.0\ \Omega$ is connected to an AC source. The source emf is given by $\mathscr{E} = \mathscr{E}_{max} \sin(\omega t + \phi_i)$, where $\mathscr{E}_{max} = 5.00$ V, $\omega = 2\pi(30\text{ Hz})$, and $\phi_i = \pi/4$. At what instants is there no current in the circuit? If you replace the resistor with a capacitor for which $C = 200$ mF, at what instants is there no current in the circuit? ••

45. In the parallel circuit of Figure P32.45, the current amplitude is the same through the inductor branch, the capacitor branch, and the resistor branch. If $L = 0.36$ mH and $C = 1.0\ \mu$F, what are the source angular frequency and the resistance R? •••

Figure P32.45

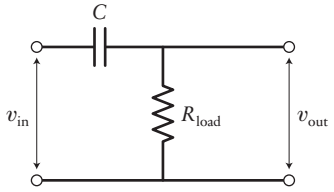


32.6 *RC* and *RLC* series circuits

46. For a given series RLC circuit, draw the variation of X_L, X_C, and impedance with frequency. ••

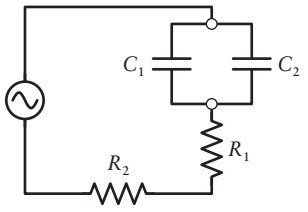
47. For a given series RLC circuit, draw the variation of phase angle with frequency. •
48. For a given series RLC circuit, draw the variation of impedance and current with frequency. •
49. A circuit consists of an AC source wired in series to a 1500-Ω resistor and a 1.5-μF capacitor. For a source emf amplitude $\mathscr{E}_{max} = 800.0$ V, calculate the phase constant ϕ and the amplitudes V_C and V_R if the source frequency is (*a*) 16 Hz and (*b*) 500 Hz. ••
50. For a series RLC circuit, charge at time t is given as $q(t) = Qe^{-Rt/2L}\cos(\omega' t + \phi)$, where Q is the amplitude at time $t = 0$. If $R = 2\ \Omega$, $C = 2.3\ \mu$F and $L = 15$ mH, find the time at which the amplitude of the charge oscillations in the circuit will be 70% of its initial value. ••
51. An AC source with $\mathscr{E}_m = 220$ V AC has a frequency of 50.0 Hz and produces the same maximum current in two series circuits. Each circuit contains a 160-Ω resistor. Circuit 1 contains a 5.6-μF capacitor, and circuit 2 contains an inductor. What is the inductance in circuit 2? ••
52. You need a filter circuit to eliminate a 60-Hz signal and lower frequencies. The load you are trying to drive has a resistance of 50 Ω. You decide to design a low-pass filter that has a cutoff frequency of 200 Hz. If you use the circuit shown in Figure P32.52, what capacitance C should you use? What fraction of any 60-Hz signal potential difference passes through this circuit? ••

Figure P32.52



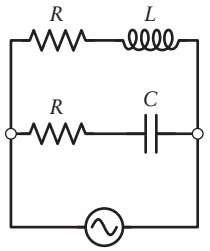
53. (*a*) Construct a phasor diagram for the circuit shown in Figure P32.53. (*b*) What is the phase constant if $C_1 = 1.00\ \mu$F, $C_2 = 3.00\ \mu$F, $R_1 = 200\ \Omega$, $R_2 = 400\ \Omega$, and $\omega = 4000\ \text{s}^{-1}$? ••

Figure P32.53



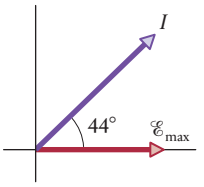
54. A loudspeaker has a resistance of 8.00 Ω. If you connect it in series with a 10.0-mH inductor, what percentage of the original maximum current are you going to get for a sound that has a pure frequency of 1000 Hz? ••
55. In the circuit shown in Figure P32.55, the current amplitude through the *RL* branch is the same as that through the *RC* branch. If $R = 10.0\ \Omega$ and $C = 50.0\ \mu$F, and the source operates at $\omega = 80.0\ \text{s}^{-1}$, what is the inductance L? ••

Figure P32.55



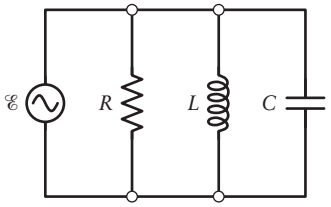
56. In problem 50, how many oscillations are completed within the time at which the amplitude of the charge oscillations in the circuit will be 40% of its initial value? ••
57. A 200-Hz power source in a series *RLC* circuit applies an AC potential difference that has an amplitude of 60 V. The current amplitude is 1.0 A, and the phasor diagram for the circuit is shown in Figure P32.57. What is the resistance of the circuit? ••

Figure P32.57



58. (*a*) What can you say, at all instants, about v_R, v_C, and v_L in the parallel circuit of Figure P32.58? (*b*) Construct a phasor diagram for V_R, V_C, V_L, and $\mathscr{E}$ and a phasor diagram for i_R, i_C, and i_L. (*c*) What is the current in the source? (*d*) Show that the current is a minimum at the frequency $\omega = \sqrt{1/LC}$. •••

Figure P32.58

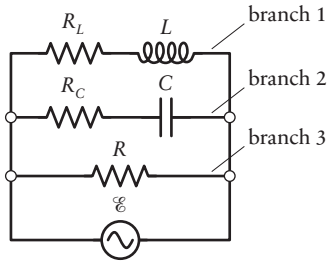


59. Two AC potential differences $v_1 = V_1\sin(\omega t + \phi_1)$ and $v_2 = V_2\sin(\omega t + \phi_2)$ are added according to

$$v_{sum} = V\sin(\omega t + \phi) = v_1 + v_2$$
$$= V_1\sin(\omega t + \phi_1) + V_2\sin(\omega t + \phi_2),$$

with $\omega = 10.0\ \text{s}^{-1}$, $V_1 = 5.00$ V, and $V_2 = 8.00$ V. At $t = 0$, $v_1(t = 0) = 4.50$ V and $v_2(t = 0) = 3.70$ V. (*a*) Use phasors to determine values for V and ϕ in the expression $v_{sum} = V\sin(\omega t + \phi)$ and a value for $v(t = 0.0500\ \text{s})$. (*b*) Repeat using trigonometry, and then compare your results from the two methods. •••
60. In the circuit shown in Figure P32.60, $\mathscr{E} = \mathscr{E}_{max}\sin\omega t$ and $\omega = 100\ \text{s}^{-1}$. If $R = 20.0\ \Omega$, $R_C = 18.0\ \Omega$, and $R_L = 15.0\ \Omega$, what must the values of L and C be in order for the current amplitude I to be the same in branches 1, 2, and 3? •••

Figure P32.60



32.7 Resonance

61. An *RLC* circuit consists of a 20-μF capacitor, a 300-Ω resistor, and a 50-mH inductor connected in series with an AC source. What is the resonant angular frequency for this circuit? •

62. A 6.00-mH solenoid is connected in series with a 1.0-μF capacitor and an AC source. The solenoid has internal resistance 3.0 Ω, which can be treated as a series resistor in the circuit. (*a*) What is the resonant angular frequency? (*b*) If the source emf amplitude is $\mathscr{E}_{max} = 15$ V, what is the current amplitude at the resonant angular frequency? •

63. A circuit contains an electromagnet that can be modeled as an inductor of inductance 0.80 H in series with a resistance of 0.50 Ω. If you want to add a capacitor in series with the electromagnet to make the circuit resonate at 380 s^{-1}, what value of *C* must you use? •

64. A series *RLC* circuit is driven by an AC source at resonance. The resistor has a resistance of $R = 10\ \Omega$, and the AC source has an emf amplitude of $\mathscr{E}_{max} = 12$ V. What is the current amplitude? •

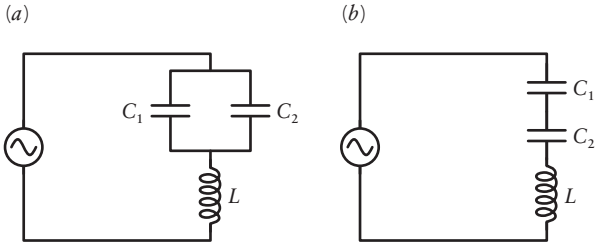
65. An *RLC* series circuit initially has resonant angular frequency ω_{0i}. Circuit elements are then changed such that both the capacitive reactance and the inductive reactance double in initial value at every operating frequency. What is the resonant angular frequency ω_{0f} after the changes are made? ••

66. A series RLC circuit consists of $R = 10\ \Omega$, $L = 0.6$ H, and $C = 10\ \mu$F. Determine the impedance at resonant frequency, 10 Hz below resonant frequency, and 10 Hz above resonant frequency. Explain the results. ••

67. An inductor resonates at 2 kHz when placed in series with a 1.80-μF capacitor. When driven at that resonant angular frequency, an AC source for which $\mathscr{E}_{max} = 180.0$ V produces a 0.3-A current in the circuit. What is the inductance? What is the resistance in the circuit? ••

68. In Figure P32.68, $C_1 = 1.00\ \mu$F, $C_2 = 2.00\ \mu$F, and $L = 47.0$ mH. Which circuit has the lower resonant angular frequency, and what is that resonant angular frequency? ••

Figure P32.68



69. An *RLC* series circuit contains an AC source of emf for which $\mathscr{E}_{max} = 30.0$ V, a 5.00-Ω resistor, a 4.00-mH inductor, and an 8.00-μF capacitor. What are (*a*) the resonant angular frequency and (*b*) the current amplitude if it is driven at the resonant angular frequency? (*c*) If a 1.00-kHz, 30.0-V source is used, what is the current amplitude? ••

70. An *RLC* series circuit has a 350-mH inductor and a 25.0-Ω resistor connected to a 50.0-Hz AC power source, and a variable capacitor. (*a*) To what capacitance value must it be set in order to have the circuit be at resonance? (*b*) If the source produces $\mathscr{E}_{rms} = 120$ V, what is the maximum potential difference across the inductor and the capacitor at resonance? ••

71. You want to build an AM radio that uses a series *RLC* circuit for tuning. The circuit consists of a 30.0-Ω resistor, a 15.0-μH inductor, and an adjustable capacitor. At what capacitance should the capacitor be set in order to receive the signal from a station that broadcasts at 870 kHz? How would you make the tuning more selective? ••

72. The resonant angular frequency is ω_0 in identical *RLC* circuits 1 and 2. Because the power delivered by the source at resonance is too high, you reduce the angular frequency of the source of circuit 1 to $\omega_1 = \omega_0/2$. To what value $\omega_2 \neq \omega_1$ must you change the angular frequency of the source of circuit 2 in order to have the power delivered by the source of circuit 1 equal that of circuit 2? •••

32.8 Power in AC circuits

73. An *RLC* circuit consists of a 220-Ω resistor, a 3.3-μF capacitor, and a 3.9-mH inductor. The circuit is connected to a power outlet that has a peak emf of 150 V and oscillates at 200 Hz. What is the average rate at which energy is dissipated in this circuit? •

74. A given capacitor of capacitance 3.5 μF is charged to 65 V. The charging battery is then disconnected and a 15-mH inductor is connected in series with this capacitor so that *LC* oscillations occur. Assume that there is no resistance in the circuit, find the maximum current in the inductor. •

75. An AC power line has a current amplitude of 4.00 A and a resistance of 0.500 Ω. At what average rate is energy lost from the line? (Assume a zero phase angle.) •

76. For household electrical outlets in the United States, the root-mean-square value of the source emf is 120 V. (*a*) For a hair dryer rated at 1875 W, what are the root-mean-square current and the current amplitude? Ignore any capacitive and inductive reactance in the hair dryer circuit. (*b*) If an equally powerful hair dryer is to be sold in Germany, where household outlets have $\mathscr{E}_{rms} = 220$ V, what are I_{rms} and I? ••

77. In the first 100.0 ms after the switch on an *RLC* circuit is closed, the resistor dissipates 5.005 J of energy. In the first 200.0 ms, the resistor dissipates 10.03 J of energy. If the circuit contains a 70.0-mH inductor, what is the smallest possible inductive reactance of the circuit? ••

78. For the circuit of Problem 69, what is the average rate at which energy is delivered by the source when the circuit is operated at the resonant frequency? When it is operated at 1.00 kHz? What is the power factor in each case? ••

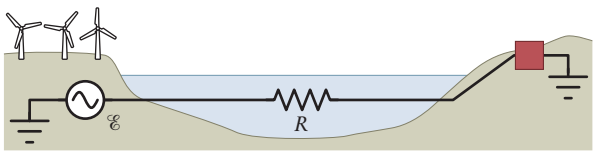
79. (*a*) Corresponding to each step in Figure 32.2 (*a*)-(*b*) [page no.1123], draw bar graphs for stored magnetic and electric energies. (*b*) What should be next time t_2 when, again, $U_E = 0.5\ U_{E,\max}$? ••

80. An AC circuit with a 240-V peak emf has a 30.0-Ω resistor connected in series to a capacitor for which the capacitive reactance is $X_C = 60.0\ \Omega$. (*a*) What is the phase constant? (*b*) What is the average rate at which energy is dissipated in the circuit? (*c*) If the resistance in the circuit is 20.0 Ω, what would the capacitive reactance have to be in order for the power factor to be 0.15? ••

81. At time $t = 0$, an LC oscillator begins to oscillate when the capacitor has its maximum charge. Assume the resistance to be zero in the circuit. If $L = 30$ mH and $C = 2.5\ \mu$F, find the first time at which the electromagnetic energy of the oscillator is shared equally between the corresponding fields of the capacitor and the inductor. ••

82. You need to design a power cable to connect an energy-storage facility located at the ocean shore to a wind turbine farm 2 km offshore (Figure P32.82). The farm can deliver energy at an average rate of 45 MW, and you need to run a single cable to deliver energy at this rate (the turbines and the storage facility share a common ground). The cable has a resistance R, leading to a loss of energy during transmission. Assume that all other loads on the circuit (electrical appliances in all the buildings serviced by this facility) have no capacitive or inductive reactance. Using different transformers at either end, you can change the value of $\mathscr{E}_{\max}$ and I while maintaining the same rate at which electrical energy is delivered to the cable. Should you choose a high or a low value for $\mathscr{E}_{\max}$? •••

Figure P32.82

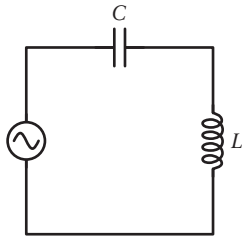


83. An *RLC* series circuit consists of a 450-Ω resistor, a 3.00-mF capacitor, and a 1.00-H inductor. The circuit is driven by a power source that oscillates at 20.0 Hz and has an $\mathscr{E}_{\text{rms}}$ value of 60.0 V. The power source is switched on at $t = 0$, and at that instant the emf is at its maximum value. Calculate the power supplied at (*a*) $t = 0.0200$ s, (*b*) $t = 0.0375$ s, and (*c*) $t = 0.500$ s. •••

Additional Problems

84. Sketch the phasor diagram for v_C, v_L, and the instantaneous current for the circuit shown in Figure P32.84 at an arbitrary instant $t \neq 0$. If the potential difference across the capacitor is a maximum at $t = 0$, what are the initial phases of v_C, v_L, and i? •

Figure P32.84



85. A capacitor is connected to an AC source. Sketch a graph showing how the current amplitude changes as you increase the AC angular frequency. •

86. Can a diode control anything other than the direction of current? ••

87. (*a*) What are the four possible combinations for input at A and B in an AND logic gate? (See Problem 30 regarding Y/N terminology for input.) What is the output for each combination? (*b*) What are the four possible combinations for input at A and B in an OR logic gate? What is the output for each combination? ••

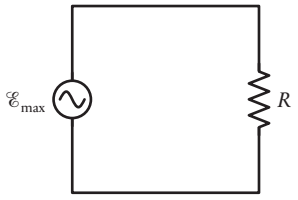
88. A solenoid can be treated as a series *RL* circuit. If such a cylindrical solenoid driven by a 60-Hz, 120-V AC source has 1200 turns of 2 cm radius wrapped around a 30-cm-long core, and a resistance of 2 Ω, what is the phase constant and the current amplitude? ••

89. Find the potential difference $v_X(t)$, its amplitude V_X, and current $i_X(t)$, its amplitude I_X ($X = R, C, L$), if you are given a (*a*) purely resistive load with resistance $R = 100\ \Omega$, (*b*) purely capacitive load with $C = 20\ \mu$F, and (*c*) purely inductive load with inductor $L = 300$ mH. In all cases, the sinusoidal alternating emf device operates at amplitude $\mathscr{E}_{\max} = 40.0$ V and frequency $f_d = 50.0$ Hz. ••

90. An *RLC* series circuit containing a 30.0-Ω resistor, a 7.5-mH inductor, and a 10-μF capacitor is driven by an AC source operating at 1200 Hz. If the source amplitude is $\mathscr{E}_{\max} = 180.0$ V, determine the amplitude of the current flowing in the circuit. ••

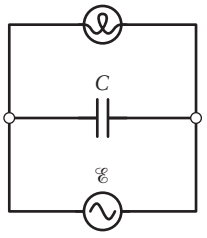
91. In the circuit shown in Figure P32.91, the AC source oscillates at 60 Hz and its emf is given by $\mathscr{E} = \mathscr{E}_{\max} \sin(2\pi ft)$, where $\mathscr{E}_{\max} = 15$ V and $R = 150\ \Omega$. Sketch, as functions of time, the current, the potential difference, and the power to the resistor. ••

Figure P32.91



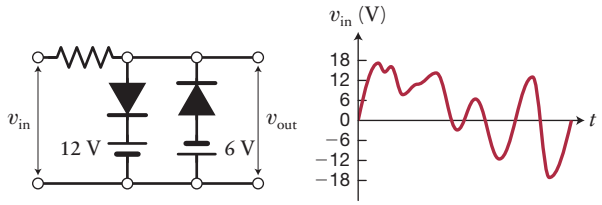
92. What happens with the light bulb shown in Figure P32.92 as you increase the frequency of the AC source? ••

Figure P32.92



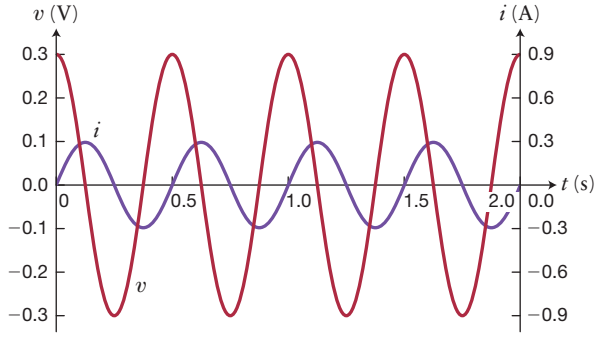
93. A diode *clipper* is a circuit used to eliminate extremes in input potential. If the graph in Figure P32.93 represents input potential as a function of time for the circuit shown, sketch a graph showing the output potential as a function of time. ●●

Figure P32.93



94. Figure P32.94 shows, for an AC circuit, the instantaneous current through a circuit element and the potential difference across the element. What can you determine about the element? ●●●

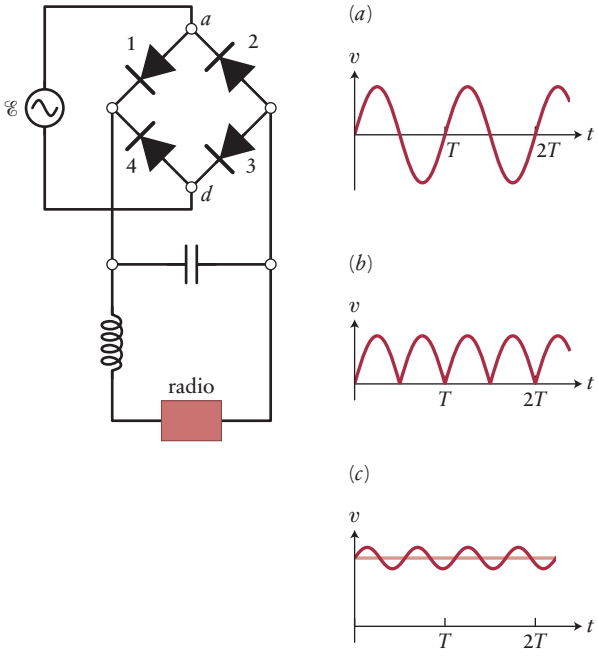
Figure P32.94



95. Your boss has purchased a new AC power source to run a high-voltage, low-current display, but it is not working. While he is fuming, you look at the owner's manual and discover that this power source produces only half the peak potential difference needed to operate the display. Because you are pretty sure you have a transformer on your workbench that will solve the problem, you tell your boss that you can fix it. When you return to your workbench, you are horrified to discover that the other members of your team have scavenged not only the transformer but most of your other supplies as well! Ignoring hookup wire and solder, you have only two identical capacitors and two identical diodes at your disposal. You collapse onto your lab stool and start thinking about how to word your apology to the boss, when suddenly you realize you can solve the problem after all! ●●● CR

96. The circuit in Figure P32.96 represents your planned design for a wall power supply that will run a radio that usually runs on a 9-V battery. The power supply uses a transformer (not shown) to convert the 110-V AC household emf of the wall outlet to a 9-V AC emf. If you do not add any additional components to this circuit, the potential difference over time is as shown in graph *a*, a typical sinusoidally varying $v(t)$ curve. Batteries supply direct current, however, and a radio designed to run on batteries cannot run on alternating current. If you address this problem by adding a rectifier to your circuit, the potential difference over time is as shown in graph *b*. The radio still will not run properly, though, because the 120-Hz cycle would either render the radio dysfunctional or result in an annoying 120-Hz buzz. By adding a large capacitor in parallel with the rectifier, you can smooth out the potential difference as the capacitor charges and discharges over the course of a cycle. Still, there is going to be a residual 120-Hz "wiggle" on top of the DC signal, as shown in graph *c*. You can filter out this wiggle by using an inductor as a low-pass filter. If the radio has an internal resistance of $R = 20.0\ \Omega$, what inductance value do you need in order to reduce the wiggle by a factor of 2? ●●●

Figure P32.96



97. The research laboratory you work in is less than a block away from the transmitting antenna of a radio station that broadcasts at 1100 kHz. You notice that every piece of equipment in the lab has 1100-kHz noise in addition to the signals you want, which are in the range from 2 kHz to 5 kHz. You wonder if there is some sort of simple circuit you could add to the lab equipment that would allow your desired frequencies to pass through without much loss, while reducing the 1100-kHz signal by at least a factor of ten. ●●● CR



33

Ray Optics

CONCEPTS

33.1 Rays
33.2 Absorption, transmission, and reflection
33.3 Refraction and dispersion
33.4 Forming images

QUANTITATIVE TOOLS

33.5 Snel's law
33.6 Thin lenses and optical instruments
33.7 Spherical mirrors
33.8 Lensmaker's formula

You can read these words because this page reflects light toward you; your eyes intercept some of the reflected light, and the lenses of your eyes redirect it, forming an image of the page on your retina. Where does the reflected light come from? Our primary source of light during the day is the Sun; our secondary source is the sky's brightness. Indoors and at night, our light sources are flames in candles, white-hot filaments or light-emitting diodes (LEDs) in light bulbs, and glowing gases in fluorescent bulbs. The light from these sources comes from the accelerated motion of electrons as this motion produces electromagnetic waves.

In Chapter 30 we studied the propagating electric and magnetic fields that constitute electromagnetic waves, and we learned that a narrow frequency range of these waves corresponds to what we know as visible light. In this chapter we continue to study light, particularly its propagation and its interactions with materials. We shall not consider the electric and magnetic fields individually, but instead think of the behavior of rays of light. Such behavior, which is called *ray optics*, was understood long before it was known that light is an electromagnetic wave.

33.1 Rays

If you pierce a small hole in a piece of cardboard and then hold the cardboard between a lamp and a screen, the position where the light transmitted through the hole strikes the screen lies on a straight line connecting the lamp and the hole (**Figure 33.1**). This observation suggests that we can think of a light source as made up of many straight beams that spread out in three dimensions from the source. Each beam travels in a straight line until it interacts with an object. That interaction changes the beam's direction of travel.

We represent the propagation of light by drawing **rays:**

A ray is a line that represents the direction in which light travels. A beam of light with a very small cross-sectional area approximately corresponds to a ray.

In order to see an object, our eyes form an image by collecting light that comes from the object. If the object is a light source, we see it by the light it emits. We can also see an object that is not a light source because such an object interacts with light that comes from a light source. The light is then redirected toward our eyes by means of this interaction.

When you stand outside on a sunny day, some of the rays from the Sun are blocked by your body while others travel in straight lines to the ground around you. You cast a shadow—a region on the ground that is darker than its surroundings because the Sun's rays that are blocked by your body do not strike this region. (The shadow region is not completely dark because it is still illuminated by light from the sky and by sunlight reflected from nearby objects.)

Figure 33.2 illustrates how rays can be used to represent the directions of light beams emanating from a light source. Just as with field line diagrams, we draw only a few rays to represent all the rays that could possibly be drawn; a ray could be drawn along any line radially outward from the source. Although most sources of light—the Sun, a flame, a light bulb—are extended, when the distance to the source is much greater than the extent of the source, we can treat that source as a *point source* (See Section 17.1). That is, we can treat the source as if all the light were emitted from a single point in space. In the first part of this chapter, we develop a feel for which rays to draw in a given situation.

Figure 33.1 A light beam that is not disturbed travels in a straight line.

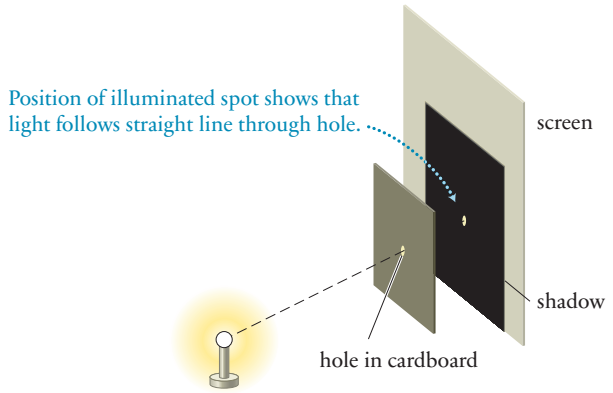
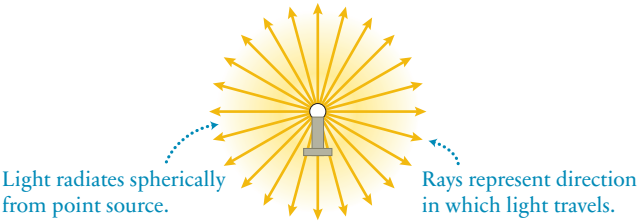


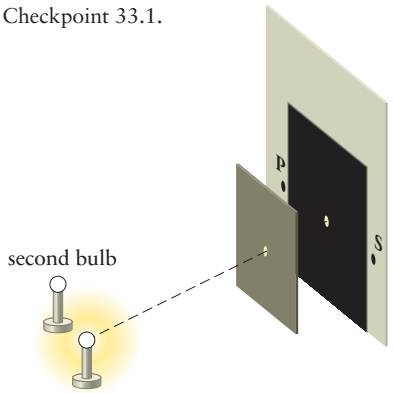
Figure 33.2 Rays emanating from a source of light.



CONCEPTS

33.1 Suppose a second bulb is added to the left of the one in Figure 33.1, as illustrated in **Figure 33.3**. What happens to (*a*) the brightness of the spot created on the screen by the first bulb and (*b*) the brightness at locations close to the vertical edges of the original shadow on the screen (points P, Q, R, and S)?

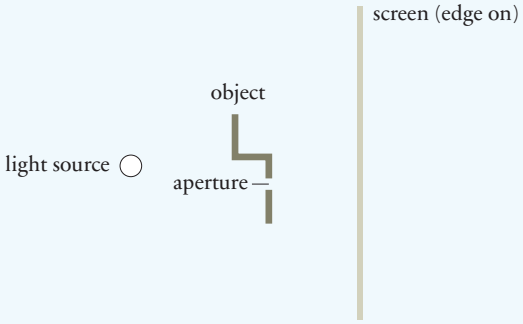
Figure 33.3 Checkpoint 33.1.



EXAMPLE 33.1 Light and shadow

An object that has a small aperture is placed between a light source and a screen, as shown in **Figure 33.4**. Which parts of the screen are in the shadow?

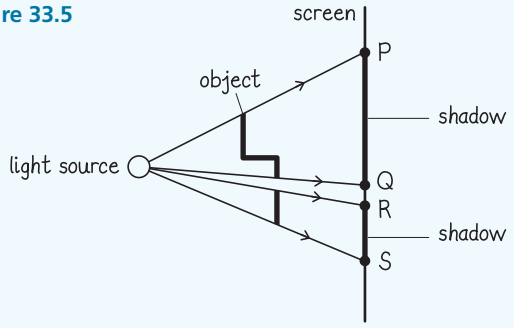
Figure 33.4 Example 33.1.



❶ **GETTING STARTED** The rays emitted by the source radiate outward in all directions following straight paths. The shadow is cast because the object prevents some of the rays from reaching the screen (except for the rays that make it through the aperture).

❷ **DEVISE PLAN** To locate the edges of the shadow, I draw straight lines from the source to the edges of the object (including the edges of the aperture) and extend these rays to the screen (**Figure 33.5**).

Figure 33.5



❸ **EXECUTE PLAN** The top and bottom edges of the shadow correspond to the highest and lowest screen locations (P and S) reached by light rays that are not blocked by the object. The gap in the shadow between locations Q and R corresponds to light rays that pass through the aperture, which means that this region of the screen is not in shadow. ✔

❹ **EVALUATE RESULT** A shadow that is taller than the object makes sense. Because the light rays from the source emanate in all directions, most of them reach the screen at angles other than 90°. This means that the distance from P to S must be greater than the object height. Indeed, I know from experience that the shadow cast by my hand gets larger as I move my hand closer to a lamp, increasing that angle.

33.2 Hold a piece of paper between your desk lamp (or any other source of light) and your desk or a wall. How does the sharpness of the edges of the shadow change as you move the paper closer to the bulb? Why does this happen?

33.2 Absorption, transmission, and reflection

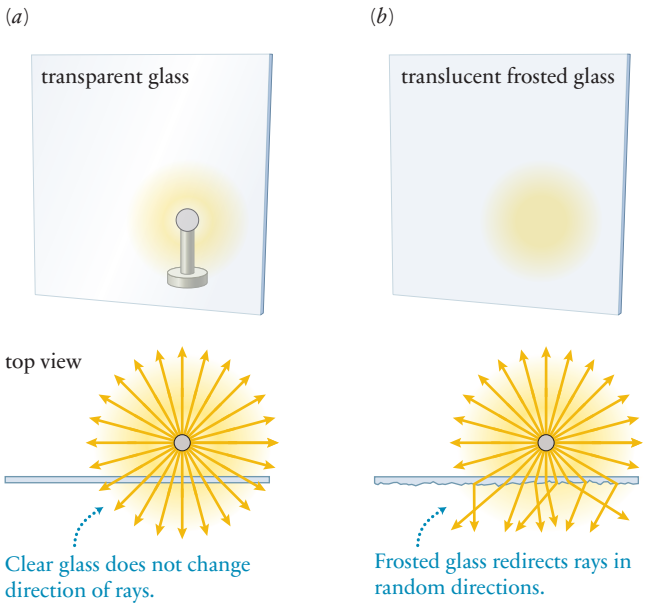
Different materials interact differently with the light that strikes them, which is how you can visually distinguish wood from metal, fabric from skin, and a white piece of paper from a blue one. When light strikes an object, the light can be transmitted, absorbed, or reflected.

Transmitted light passes through a material. Objects that transmit light, such as a piece of glass, are said to be *transparent* (**Figure 33.6a**). In *translucent* materials, such as frosted glass, light rays are *transmitted diffusely*—that is, they are redirected in random directions as they pass through, so that the transmitted light does not come from a definite direction (Figure 33.6*b*). Because translucent materials scatter light in this manner, we cannot see objects clearly through them.

Absorbed light enters a material but never exits again. Objects that absorb most of the light that strikes them, such as a piece of wood, are said to be *opaque*. When light strikes such materials, the energy carried by the light is converted to some other form (usually thermal energy) and the light propagation stops.

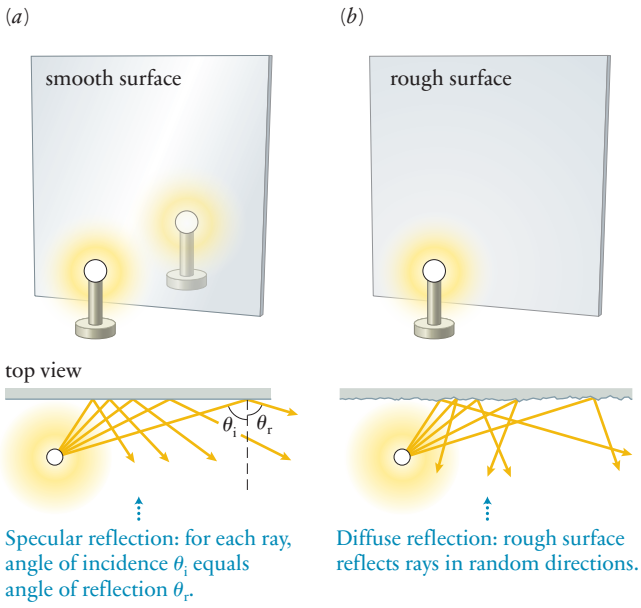
Reflected light is any light that is redirected away from the surface of the material (**Figure 33.7**). Smooth surfaces reflect light *specularly*—that is, each ray bounces off the surface in such a way that the angle between it and the normal to the surface doesn't change (Figure 33.7*a*). The angle between the incoming ray and the normal to the surface is called the **angle of incidence** θ_i; the angle between the outgoing ray and the normal is called the **angle of reflection** θ_r.

Figure 33.6 We see objects clearly through a sheet of clear glass but diffusely through frosted glass.



CONCEPTS

Figure 33.7 Light reflects specularly from a smooth surface, forming a mirror image. From a rough surface, it reflects diffusely (in random directions), so no image forms.



Empirically we find:

For a ray striking a smooth surface, the angle of reflection is equal to the angle of incidence, and both angles are in the same plane.

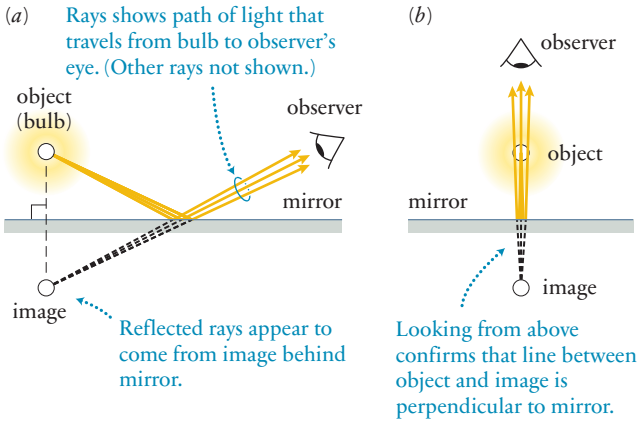
This **law of reflection** holds at smooth surfaces for any angle of incidence.

Surfaces that are not smooth reflect light in many directions (Figure 33.7*b*). For such *diffuse reflection*, each ray obeys the law of reflection, but the direction of the surface normal varies over the surface and so the angle of reflection also varies.

How smooth is smooth? If the height and separation of irregularities on the surface are small relative to the wavelength of the incident light, the surface acts like a smooth surface and most light is reflected specularly. For example, paper appears smooth to microwaves, which have wavelengths ranging from millimeters to meters, and therefore microwaves are reflected specularly from paper. Visible light, however, has wavelengths of hundreds of nanometers, and so paper reflects visible light diffusely.

Rays that come from an object and are reflected from a smooth surface form an **image,** an optically formed duplicate of the object (Figure 33.7*a*). **Figure 33.8***a* shows the paths taken by light rays emitted by a light bulb placed in front of a mirror. A diagram like Figure 33.8*a* showing just a few selected rays is called a **ray diagram.** If we trace the reflected rays back to the point at which they appear to intersect, we see that point is behind the mirror. Consequently, the brain interprets the reflected rays as having come from that point, creating the illusion that the light bulb is behind the mirror. The directions of the rays that reach the eyes of the observer are the same as if they had come from an object located behind the mirror.

Figure 33.8 Diagrams showing the paths taken by light rays that are produced by a bulb and reflected by a mirror into an observer's eye. The reflected rays appear to come from behind the mirror, forming an image behind the mirror.



Note that the image is located on the line through the object and perpendicular to the mirror. If we look along that line, the image lies behind the object (Figure 33.8*b*).

Rays that do not actually travel through the point from which they appear to come, like the rays in Figure 33.8*a*, are said to form a *virtual image*. A *real image* is formed when the rays actually do intersect at the location of the image. (Flat mirrors cannot form real images; we'll encounter real images when we discuss lenses in Section 33.4 and curved mirrors in Section 33.7.)

EXAMPLE 33.2 How far behind the mirror?

If the light bulb in Figure 33.8*a* is 1.0 m in front of the mirror, how far behind the mirror is the image?

❶ **GETTING STARTED** The location of the image is the location from which the rays reflected by the mirror appear to come—that is, the point at which they intersect. From Figure 33.8 I know that because the rays intersect directly behind the bulb, a line that passes through the bulb and is normal to the mirror passes through the image.

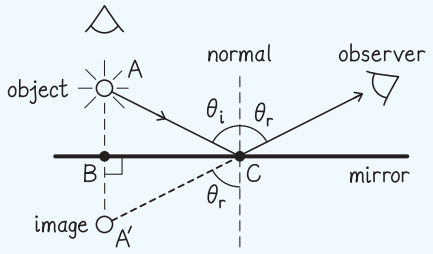
❷ **DEVISE PLAN** I can obtain the distance of the image behind the mirror by considering one ray that travels from the bulb to the observer and then tracing that ray back through the mirror to its intersection with the line that is perpendicular to the mirror and passes through the bulb.

❸ **EXECUTE PLAN** I begin by drawing a ray that travels from the bulb to the mirror and is reflected to the location of the observer (**Figure 33.9** on the next page). In my drawing, A denotes the bulb location, B denotes the point where the line connecting the bulb and its image intersects the mirror, and C denotes the point at which the ray that is reflected to the observer hits the mirror. According to the law of reflection, $\theta_r = \theta_i$.

(Continued)

CONCEPTS

Figure 33.9



I now extend the reflected ray to behind the mirror (dashed line). I know that the image must lie somewhere along that dashed line and must also lie on the line that passes through the object and is perpendicular to the mirror. The image must therefore lie at the intersection of this line and the dashed ray extension; I denote that intersection point by A′.

To determine the distance BA′, which is how far behind the mirror the image is, I note that angle A′CB is equal to $90° - \theta_r$ and angle ACB is equal to $90° - \theta_i$. Because $\theta_i = \theta_r$, angles A′CB and ACB are equal. Therefore triangles ABC and A′BC are congruent, and AB = BA′. That is, the image appears at the same distance behind the mirror as the object is in front of it: 1.0 m behind the mirror. ✔

4 EVALUATE RESULT My result makes sense because I know from experience that as I walk toward a mirror, my image also approaches it.

33.3 If the observer in Figure 33.8 moves to a different position, does the location of the image change?

The colors of visible light we see correspond to different frequencies of electromagnetic waves. Red corresponds to the lowest frequency of the visible spectrum. As the frequency increases, the color changes to orange, yellow, green, blue, indigo, and finally violet (the highest frequency of the visible spectrum). The range of visible frequencies is quite small relative to the range of the complete electromagnetic spectrum, as **Figure 33.10** shows. Frequencies lower than the visible correspond to infrared radiation, and higher frequencies to ultraviolet. When a light source produces all the frequencies

Figure 33.10 Visible light makes up only a small part of the electromagnetic spectrum.

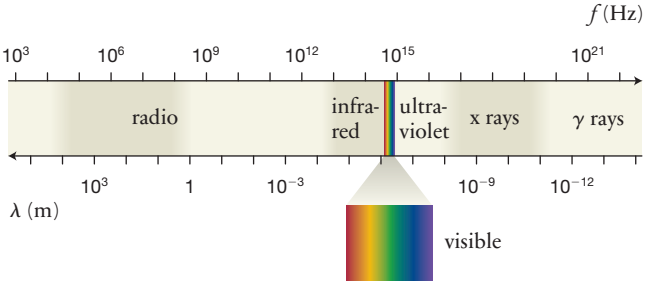
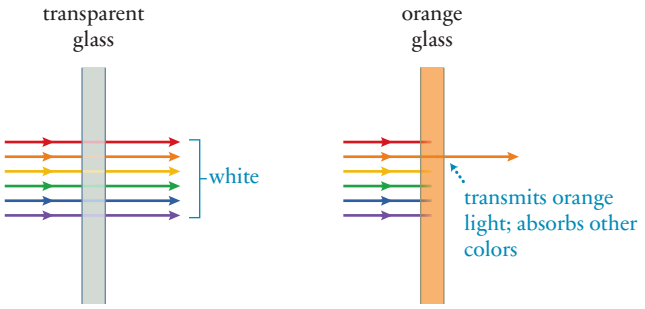


Figure 33.11 All colors of light pass through colorless glass (shown gray for illustration purposes); orange glass transmits orange light and absorbs all colors of light except orange.

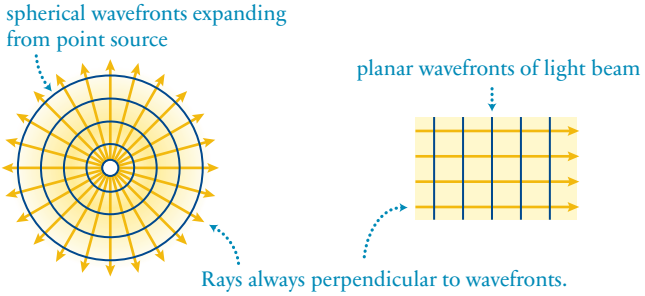


of the visible spectrum at roughly the same intensities, the emitted light appears white.

Different colors of light interact differently with different objects, affecting the color we perceive the object as being. Colorless materials, like a piece of ordinary window glass, transmit all colors of the visible spectrum. A piece of orange glass, on the other hand, transmits only the orange part of the visible spectrum. All other colors are absorbed in the glass (**Figure 33.11**). A red apple absorbs all colors of the visible spectrum except red, which is redirected to our eyes. Grass absorbs all colors except green, which is diffusively reflected at its surface.

Because light is a wave phenomenon, it is sometimes useful to represent the propagation of light with wavefronts, which we introduced in Section 17.1. Wavefronts are drawn perpendicular to the direction of propagation of the wave.* Because light rays point along the direction of propagation of the light, light wavefronts are perpendicular to light rays. **Figure 33.12*a*** shows the spherical wavefronts for light coming from a point source, and Figure 33.12*b* shows the straight-line rays and wavefronts corresponding to a planar electromagnetic wave.

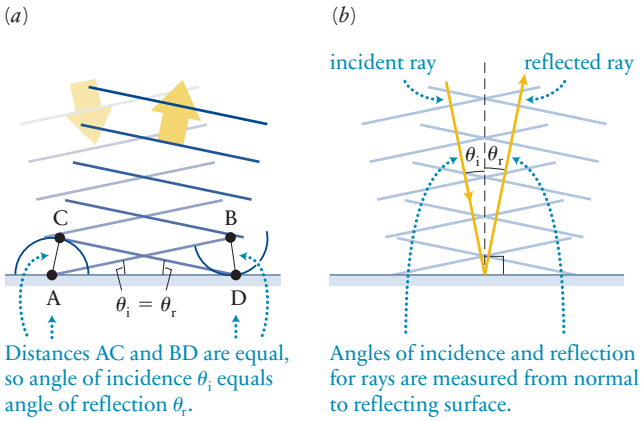
Figure 33.12 A point source of light produces spherical wavefronts; a beam of light contains planar wavefronts.



*For mechanical waves, the wavefronts are drawn at the locations of the wave crests, spaced by the wavelength of the wave. Because the wavelength of light is very short, the wavefronts of light cannot be represented to scale.

CONCEPTS

Figure 33.13 (*a*) The reflection of wavefronts from a smooth surface explains the law of reflection. (*b*) The corresponding rays and their angles of incidence θ_i and reflection θ_r.

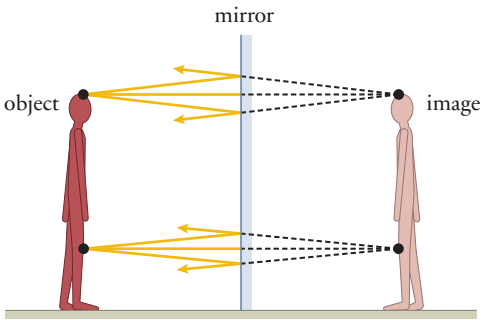


Note that a planar wave is represented with rays that are parallel to one another because all the wavefronts are parallel to one another.

By looking at how wavefronts behave, we can understand the law of reflection. When a light ray strikes a smooth surface at an incidence angle $\theta_i \neq 0$ (**Figure 33.13**), the left end of the first wavefront to reach the surface gets there, at A, before the right end does. In the time interval it takes the right end to reach the surface at D, the left end has traveled back from the surface to C. The distance traveled by the right end toward the surface, BD, is the same as that traveled by the left end away from the surface, AC, so the angles BAD and CDA must be equal. The angle of incidence θ_i equals angle BAD. Likewise, the angle of reflection θ_r equals angle CDA. So $\theta_i = \theta_r$.

So far we have treated the object (and consequently the image) as a single point. **Figure 33.14** shows how images are formed of extended objects. Each point on the object reflects (or emits) light rays, and the reflections of these rays appear to come from a corresponding point on the image. A flat mirror thus produces behind the mirror an exact mirror image of the entire extended object.

Figure 33.14 Paths taken by rays from more than one point on the object, showing how extended images form.



33.4 In order for the person in Figure 33.14 to see a complete image of himself, does the mirror need to be as tall as he is?

33.3 Refraction and dispersion

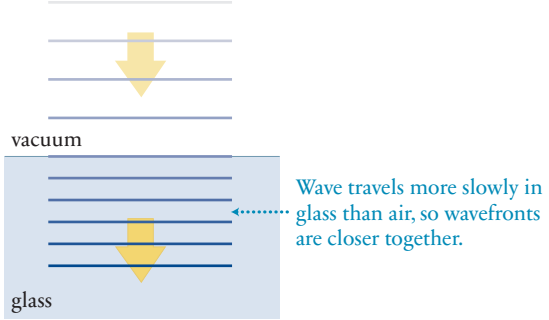
As we found in Chapter 30, light propagates with speed $c_0 = 3 \times 10^8$ m/s in vacuum. In air, the speed of light is almost the same as that in vacuum. In a solid or liquid medium, however, light propagates at a speed c that is generally less than c_0. In glass, for example, visible light propagates at two-thirds of the speed of light in vacuum (see Example 30.8).

How does this change in speed affect the propagation of an electromagnetic wave? Recall from Chapter 16 that harmonic waves are characterized by both a wavelength λ and a frequency f and that the product of the wavelength and frequency equals the wave's speed of propagation (Eq. 16.10). The frequency of the wave must remain the same because the oscillation frequency of the electromagnetic field that makes up the wave is determined by the acceleration of charged particles at the wave's source. The acceleration of the source does not alter when the wave travels from one medium to another, and thus the frequency of the traveling wave also cannot change.

33.5 In vacuum, a particular light wave has a wavelength of 400 nm. It then travels into a piece of glass, where its speed decreases to two-thirds of its vacuum speed. What is the distance between the wavefronts in the glass?

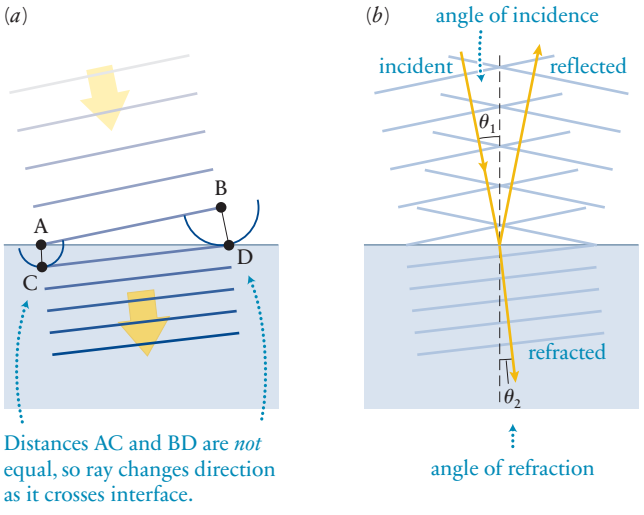
As we found in Checkpoint 33.5, when rays of light pass through the interface between vacuum and a transparent material, the wavefronts inside the material are more closely spaced than they are in vacuum, due to the lower speed of the wavefronts. **Figure 33.15** illustrates this effect for wavefronts incident normal to the surface of the material.

Figure 33.15 Wavefronts for a ray traveling from vacuum into transparent glass in a direction normal to the glass surface.



CONCEPTS

Figure 33.16 (*a*) Refraction is explained by the behavior of wavefronts that cross at an angle into a transparent medium in which they travel more slowly. (*b*) Incident, reflected, and refracted rays, showing the angles of incidence θ_1 and refraction θ_2 (measured from the normal to the surface).



When a light ray travels from one material into a second material where light travels more slowly, the ray bends toward the normal to the interface between the materials.

Generally, the speed of light decreases as the mass density of the material increases. Note also that, as shown in Figure 33.16*b*, both reflection and refraction take place at the interface between two media (or between vacuum and a medium).

The amount of bending depends on the angle of incidence and on the relative speeds in the two media. There is no bending for normal incidence (as we saw in Figure 33.15); the bending is less near normal incidence and becomes more pronounced as the angle of incidence increases. In Section 33.5, we'll work out a quantitative expression relating angles θ_1 and θ_2.

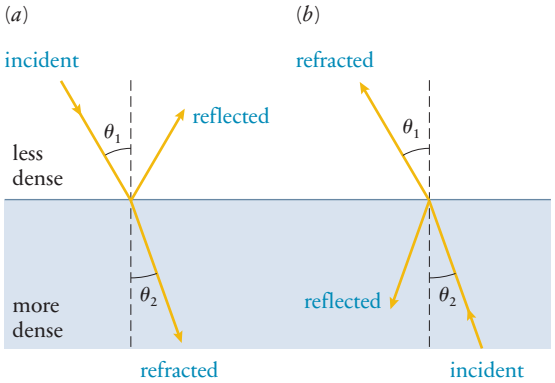
33.6 Suppose the ray in Figure 33.16 travels in the opposite direction—that is, from the more dense medium to the less dense medium. If the angle of incidence is now θ_2, how does the angle of refraction compare with θ_1?

What if the wavefronts strike the transparent material at an angle? In such a case, one end of the wavefront arrives at the surface before the other (**Figure 33.16**). Once the end that reaches the surface first (this happens to be the left end in Figure 33.16) enters the material, it travels at the lower speed while the other end of the wavefront (the right end in our example) continues to travel at the vacuum speed. This means the distance AC traveled by the left end is less than the distance BD traveled by the right end during the same time interval (Figure 33.16*a*). Consequently, the wavefront CD in the material is no longer parallel to the wavefront AB that has not yet entered the material.

The direction of the ray associated with these wavefronts therefore changes on entering the material. As shown in Figure 33.16*b*, the angle of incidence θ_1, between the ray in vacuum and the normal to the interface between the two materials, is greater than the angle θ_2, between the ray in the material and the normal to that interface. This bending of light as it moves from one material into another is referred to as **refraction,** and the angle θ_2 between the refracted ray and the normal to the interface between the materials is called the **angle of refraction.** Whenever light is refracted, the angle between the ray and the normal is always greater in the material in which the light travels faster, so:

Because the relationship between the angles of incidence and refraction is completely determined by the speed of the wavefronts in the two media, the angles do not depend on which is the incident ray and which is the refracted ray. As shown in **Figure 33.17**, θ_1 and θ_2 have the same values whether θ_1 is the angle of incidence (Figure 33.17*a*) or the angle of refraction (Figure 33.17*b*). Keep in mind, however, that the reflected ray is always on the same side of the interface as the incident ray, and so the angle of reflection is *not* the same in Figure 33.17*a* and Figure 33.17*b*.

Figure 33.17 Because refraction is caused by the relative speeds of the wavefronts in two media, the angles of incidence and refraction do not depend on which is the incident ray and which the refracted ray.



CONCEPTS

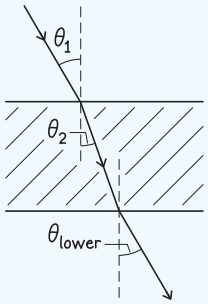
EXAMPLE 33.3 Crossing a slab

Consider a light ray incident on a parallel-sided slab of glass surrounded by air, as shown in Figure 33.17*a*. The ray travels all the way through the slab and emerges into air on the other side. In what direction does the ray emerge?

❶ **GETTING STARTED** This problem involves two successive encounters of a light ray with interfaces between glass and air. At each interface, the ray is refracted. I need to determine the direction of the ray (its angle to the normal to the slab) after it crosses the lower interface of the slab represented in Figure 33.17*a*.

❷ **DEVISE PLAN** Because I want to know the direction of the emerging ray, I construct an appropriate ray diagram. Figure 33.17*a* shows the direction of the ray inside the slab. I extend the ray through the slab to the lower interface (**Figure 33.18**) and draw the emerging ray, labeling its angle to the normal θ_{lower}. To determine this angle, I need to consider the refraction that occurs at the lower interface.

Figure 33.18



❸ **EXECUTE PLAN** Because the two interfaces are parallel, their normals are also parallel, and so the angle at which the ray is incident on the lower interface is equal to the angle θ_2 at which it is refracted at the upper interface. I saw in Figure 33.17*b* that if the angle between the ray and the normal in the slab is θ_2, it doesn't matter whether the ray in the slab is the incident ray or the refracted ray; either way, the angle between the ray in the air and the normal is θ_1. Therefore $\theta_{\text{lower}} = \theta_1$. ✔

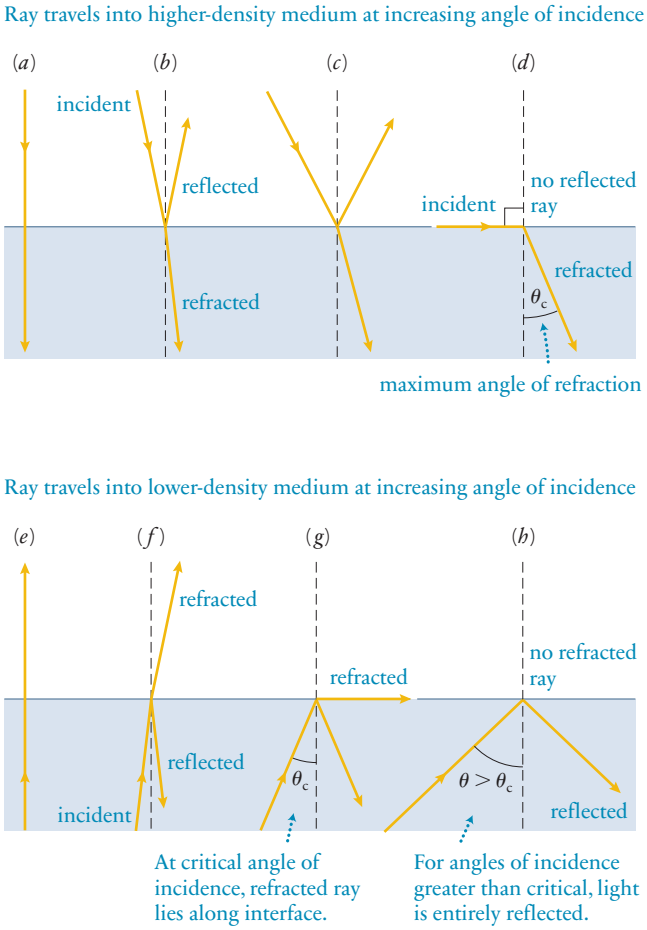
❹ **EVALUATE RESULT** Crossing the lower interface from glass into air, the ray bends away from the normal, as it should because glass is denser than air. With $\theta_{\text{lower}} = \theta_1$, in fact, the ray emerges parallel to the original direction it had before entering the slab. This makes sense because the two air-glass interfaces are parallel. (Note, though, that the ray is shifted sideways by a small distance relative to its original path.)



33.7 When the ray reflected from the bottom surface in Figure 33.18 reemerges from the top surface, how does the angle it makes with the normal compare with θ_1?

What range of refraction angles is possible? To answer this question, let's first consider the case where the ray travels from a low-density medium into a denser medium (**Figures 33.19*a–d***). Because the angle of incidence is always greater than the angle of refraction in this situation, as the angle of incidence approaches 90°, the angle of refraction remains less than 90° (Figure 33.19*d*). The full 90° range of incidence angles gives a range of refraction angles that is less than 90°.

Figure 33.19 The range of possible refraction angles for a ray crossing into a medium of either higher or lower density.



Next consider the case where the ray travels from a high-density medium into a lower-density medium (Figures 34.19*e–g*). The angle of incidence is now less than the angle of refraction. Consequently, as the angle of incidence increases, it reaches a value for which the refracted ray emerges along the interface (Figure 33.19*g*). This angle of incidence is called the **critical angle** θ_c and is equal to the angle of refraction shown in Figure 33.19*d*. For angles of incidence greater than θ_c, the angle of refraction would have to be greater than 90°, which is impossible. Therefore, no light is refracted. Instead, all the light is *reflected* back into the higher-density medium (Figure 33.19*h*), a phenomenon called **total internal reflection.**

Several optical devices make use of total internal reflection to direct light. The glass prism shown in

CONCEPTS

Figure 33.20 A prism can act as a perfect mirror by means of total internal reflection.

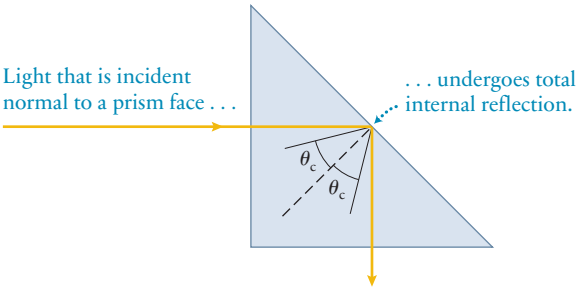
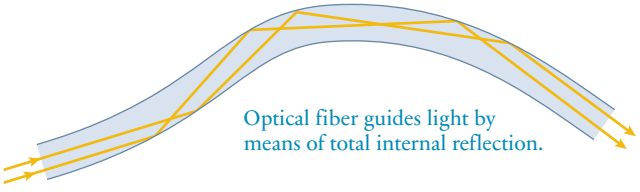


Figure 33.20 reflects light just as a mirror would. A light ray enters the prism's front surface at normal incidence. Because the back surface is slanted relative to the front surface, the angle at which the ray hits the back surface is less than 90°. This back-surface angle of incidence is greater than the critical angle for the glass, however, and so the light is totally reflected from the back surface. Such prisms are actually better mirrors than most regular mirrors; they reflect very close to 100% of the incident light, whereas mirrors are less reflective due to imperfections in the reflecting surface.

Optical fibers also guide light by means of total internal reflection. An optical fiber is a long, thin fiber made of a transparent material such as glass. If light shines into one end of the fiber at an angle greater than the critical angle, the light travels along the fiber through repeated total internal reflections, and essentially all of the light that entered the fiber emerges at the other end (**Figure 33.21**). Because very little light is lost as the light travels, only a faint glow comes from the rest of the fiber.

CONCEPTS

Figure 33.21 How optical fibers work.



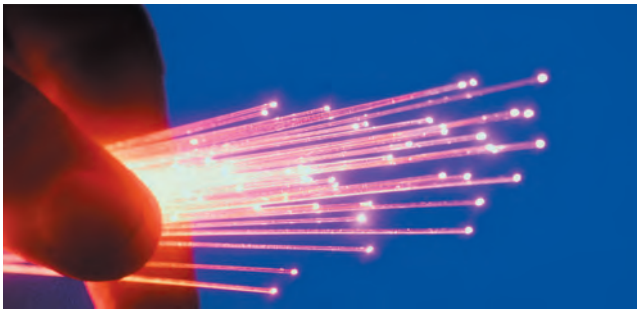
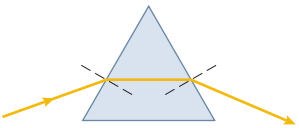
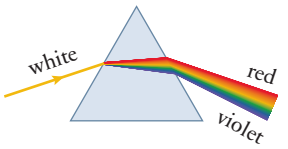


Figure 33.22 The phenomenon of dispersion, which results from the fact that the speed of light in a given medium (and hence the angle of refraction) depends slightly on the frequency of the light.

(*a*) Prism refracts light of single frequency



(*b*) Dispersion: different colors have different angles of refraction



(*c*) Rainbows result from dispersion of sunlight by raindrops



Because the speed of light in any given medium depends slightly on the frequency of the light, the angle of refraction also depends on frequency. This phenomenon is called **dispersion** because it causes rays of different colors to separate—to be *dispersed*—when refracted. Prisms like the one shown in **Figure 33.22** are designed to separate colors by the frequency dependence of the angle of refraction. In most media, high-frequency light travels more slowly than low-frequency light, and so high-frequency light bends more strongly toward the normal. The lowest frequency of visible light is red and the highest is violet, which means violet light bends the most, as the rainbow of Figure 33.22*c* shows.

Both rainbows and the brilliance of gems result from a combination of total internal reflection and dispersion. In a rainbow (Figure 33.22*c*), the combination of total internal reflection and dispersion means that we see different colors coming from water droplets at different viewing angles. Gems are cut with many internal surfaces from which total internal reflection takes place. Because the light is also dispersed, colorless gems such as diamonds shine with many distinct colors.

33.8 Because of dispersion, the critical angle for total internal reflection in a given medium varies with frequency. Is the critical angle for a violet ray greater or less than that for a red ray?

Fermat's principle

Figure 33.23 shows four ways in which a light ray can travel between two locations A and B: directly, reflected from a mirror, refracted through a glass slab, and refracted through a prism.* You could say that in each case the ray reaches B because it is aimed properly from A. However, an entirely different way of looking at the path followed by the light was suggested by the French mathematician Pierre de Fermat (1601–1665) in a formulation today known as **Fermat's principle:**

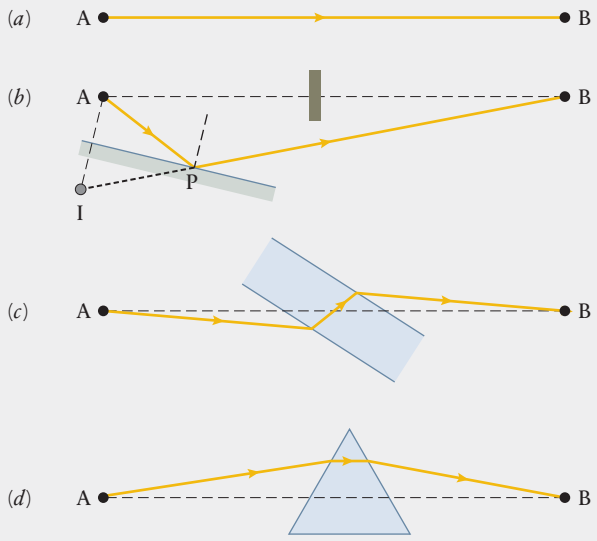
> **The path taken by a light ray between two locations is the path for which the time interval needed to travel between those locations is a minimum.**

This principle may seem to imply that light always travels in a straight line. However, the *quickest path* between two locations is not necessarily the *shortest distance* when the speed of light differs in different regions.

Let's consider the four paths in Figure 33.23 using Fermat's principle. In Figure 33.23*a*, the ray does follow a straight path because the medium in which the ray travels is uniform. As a result, the quickest path is indeed the shortest distance: a straight line from A to B.

In Figure 33.23*b*, the fact that the straight-line path from A to B is blocked means that the ray must reflect somewhere off the mirror in order to travel from A to B. The path shown, which satisfies the law of reflection, is the shortest distance from A to B involving reflection from the mirror. Because the distance from A to the reflection location P equals the distance between the image location I and P, the straight line IB is equal in length to the path traveled by the ray from A to B. Moving the reflection location to either side of P, so that the angle of incidence does not equal the angle of reflection, increases the length of the path. Thus, Fermat's principle implies the law of reflection.

Figure 33.23 Ray diagrams illustrating the *quickest path* for a light ray traveling from A to B for four situations.



When the ray must travel through some air and some glass, as in Figures 33.22*c* and *d*, the quickest path is not a straight line because the ray's speed in the glass is only two-thirds of its speed in air. To minimize the time interval needed to travel from A to B, the ray bends on entering and exiting the glass. Such a bent path reduces the distance traveled through the glass without increasing the distance traveled in air so much that it offsets the amount of time saved. In Example 33.7, we shall see that calculating the bending angles with Fermat's principle gives the same result as with ray optics.

*Note that what distinguishes a glass slab from a glass prism is the way I use the terms: In a slab, the two opposite surfaces are parallel to each other; in a prism, they are not.

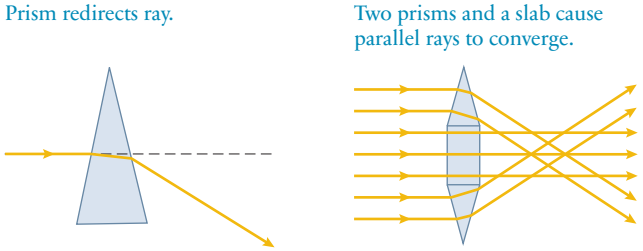
33.4 Forming images

As shown in **Figure 33.24**, by combining two prisms and a glass slab we can create a device that steers parallel light rays toward each other. The rays through the center of the device pass straight through, those through the top prism are refracted downward, and those through the bottom prism are refracted upward.

To bring all parallel incident rays to a single point, a structure called a **lens** is used. A lens is designed with curved surfaces so that the refraction of incident rays increases gradually as we move away from the center. To accomplish this, lenses are typically made with spherical surfaces, which are easy to manufacture.

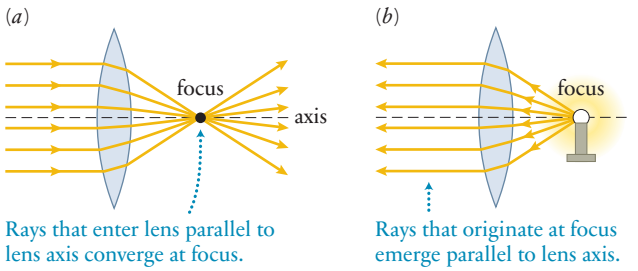
Figure 33.25*a* on the next page is a cross section of a lens with *convex* spherical surfaces, where a *convex surface* is defined as one that curves like the outside of a sphere. Rays parallel to the lens *axis*—a line perpendicular to the lens through its center—converge through such a lens onto a single point called either the **focus** or the **focal point.** A lens with convex surfaces is therefore called a *converging lens*. The distance from the center of the lens to the focus is called the **focal length** *f*.

Figure 33.24 A device that redirects parallel light rays toward each other.



CONCEPTS

Figure 33.25 Converging lens with convex spherical surfaces.



What if we place a light source at the focus of a lens, as in Figure 33.25*b*? As we saw in the preceding section, the path followed by a light ray is unaffected by reversing the direction of propagation of the ray, as long as the ray is not absorbed by the medium. So, if we place a light source at the focus of a lens, a beam of parallel rays emerges on the other side of the lens.

33.9 Sketch the wavefronts corresponding to all the rays in Figure 33.25*a*, both the parallel ones on the left and the refracted ones on the right.

We can reverse the direction of the rays through the lens of Figure 33.25, so that parallel rays enter from the right (**Figure 33.26**). The rays converge again at a focus that is the same distance f from the center of the lens. Thus, every lens has two foci, one on either side of the lens at the same distance f from it.

If we tilt the parallel rays a bit relative to the lens axis (**Figure 33.27**), the rays still converge at a distance f from the center, but the focus is no longer on the axis. Provided the parallel rays make only a small angle with the lens axis, they all converge at a point on a plane—called the *focal plane*—that is perpendicular to the axis a distance f from the lens. Rays that run near the lens axis—either parallel to it or at a small angle—are said to be *paraxial*.

Now that we know how parallel rays and rays that emanate from the focus of a lens are refracted by the lens, we can determine where images are formed. The image of a point on an object is formed where all the light rays emanating from that point converge. (These light rays then diverge from the location of the image; when they enter the eye, the brain interprets them as having come from the location of the image.)

Figure 33.26 A lens has two equivalent foci, one on each side, at equal distances from the center of the lens.

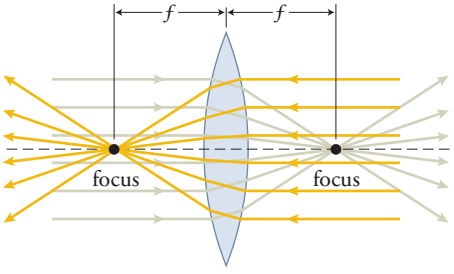
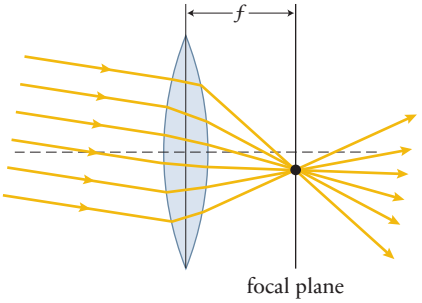


Figure 33.27 If the rays strike the lens at an angle, they no longer converge on the focus, but they still converge on a focal plane at the focal distance f.



An image of the entire object is made up of the images of all the individual points on the object.

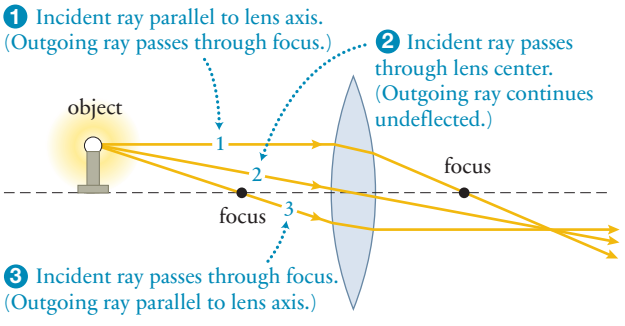
To determine where the rays emanating from a point on an object converge, we don't need to draw all the rays. Instead, we draw three special ones, called **principal rays,** and see where they converge:

1. a ray that travels parallel to the lens axis before entering the lens,
2. a ray that passes through the center of the lens, and
3. a ray that passes through the focus that is on the same side of the lens as the object.

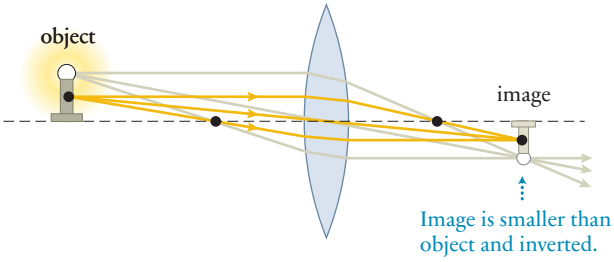
These three principal rays are shown in **Figure 33.28*a*** for the case where the object lies beyond the focus of the lens. We already know how rays 1 and 3 travel. Ray 1

Figure 33.28 The three principal rays for a spherical lens can be used to determine the location, size, and orientation of the image for a given object.

(*a*) The three principal rays



(*b*) Using principal rays to determine location and orientation of image



CONCEPTS

PROCEDURE: Simplified ray diagrams for lenses

To determine the location and orientation of an image formed by a lens, follow this procedure:

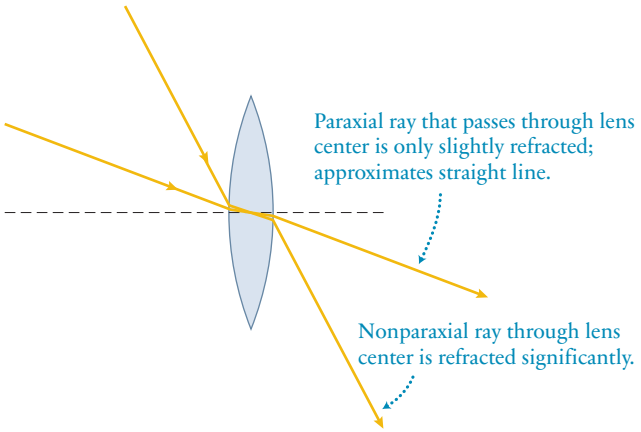
1. Draw a horizontal line representing the lens axis (the line perpendicular to the lens through its center). In the center of the diagram, draw a vertical line representing the lens. Put a + above the line to represent a converging lens or a − to represent a diverging lens.
2. Put two dots on the axis on either side of the lens to represent the foci of the lens. The dots should be equidistant from the lens.
3. Represent the object by drawing an upward-pointing arrow from the axis at the appropriate relative distance from the lens. For example, if the distance from the object to the lens is twice the focal length of the lens, put the arrow twice as far from the lens as the dot you drew in step 2. The top of the arrow should be at about half the height of the lens.
4. From the top of the arrow representing the object draw two or three of the three *principal rays* listed in the Procedure box "Principal rays for lenses" on page 1177.
5. The top of the image is at the point where the rays *that exit the lens* intersect (if they diverge, trace them backward to determine the point of intersection). If the intersection is on the opposite side of the lens from the object, the image is real; if it is on the same side, the image is virtual. Draw an arrow pointing from the axis to the intersection to represent the image (use a dashed arrow for a virtual image).

In general it is sufficient to draw two principal rays, but depending on the situation, some rays may be easier to draw than others. You can also use a third ray to verify that it, too, goes through the intersection. (If it doesn't, you have made a mistake.)

passes through the focus on the other side of the lens, and ray 3 emerges from the lens parallel to the axis. As for ray 2, as long as it is paraxial, it passes straight through with negligible refraction (**Figure 33.29**). (Nonparaxial rays are shifted significantly; our treatment of lenses in this chapter is restricted to images formed by paraxial rays.) Ray 2 can therefore be drawn as traveling in a straight line through the center of the lens.

To determine the location and orientation of the image of an extended object, we work out the locations of the images of several points on the object. Figure 33.28*b* shows where the rays converge for two points on the object, and the extended image that can be inferred from these points. The image is smaller than the object and inverted. See the Procedure box "Simplified ray diagrams for lenses" above for a general description of how to draw simplified ray diagrams for lenses.

Figure 33.29 Paraxial and nonparaxial rays. For a paraxial ray or a lens that is not too thick, the refraction is so slight that we can consider this ray to be one uninterrupted straight line.



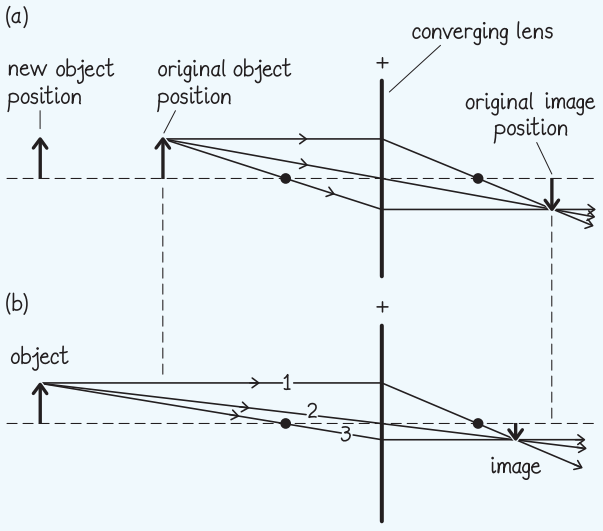
33.10 Do you need to draw all three principal rays to determine the location of an image?

EXAMPLE 33.4 Where is the image?

Consider the light bulb that is the object in Figure 33.28. If you move the bulb to the left, does the image shift left, shift right, or stay in the same place?

❶ **GETTING STARTED** Using the procedure for drawing simplified ray diagrams, I represent the lens, object, and image of Figure 33.28, and draw the bulb at its new position (**Figure 33.30*a***).

Figure 33.30



(Continued)

CONCEPTS

2 DEVISE PLAN I can determine which way the image shifts by drawing the principal rays for the light bulb in its new position.

3 EXECUTE PLAN With the rays drawn (Figure 33.30*b*), I see that the image shifts left. ✔

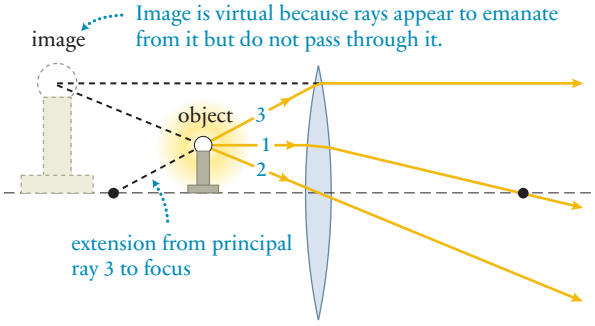
4 EVALUATE RESULT As I move the light bulb to the left, principal ray 1 remains the same, but principal ray 2 makes less of an angle with the lens axis than before. Consequently, the location at which these two rays intersect is closer to the lens than before, shifting the image to the left, as I concluded from the diagram.

Notice in Figure 33.30 that the image gets smaller as the object moves farther away from the lens. Conversely, moving the object closer to the lens makes the image larger. Indeed, one of the most common uses of lenses is to enlarge images. If the object is placed at the focus, no image forms because the rays all emerge from the lens parallel to each other—in other words, the rays do not converge.

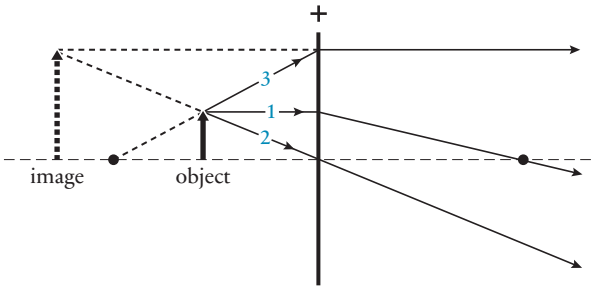
Consider placing an object *between* the lens and its focus, as in **Figure 33.31**. In this configuration, principal ray 3 does not pass through the focus. Instead, it lies on the line that joins the focus to the point of interest on the object.

Figure 33.31 (*a*) When an object is located between the lens and the focus, the image is virtual and enlarged. (*b*) In a simplified ray diagram, the object and image are replaced with solid and dashed arrows, respectively, and the lens is replaced by a vertical line. (The + indicates a converging lens.)

(*a*) Ray diagram for an object located between the focus and the lens



(*b*) Simplified version of ray diagram



The image formed in the configuration of Figure 33.28 (object beyond focus) is real because the rays really do converge at the point where the image is formed. In contrast, the image in Figure 33.31 is virtual because the rays do not actually converge at the point where the image is formed. (The extensions of these rays do cross the image point, however, and so an observer interprets the rays emerging from the lens as having traveled along straight lines from the location of the image, as indicated in Figure 33.31*a*.)

An important difference between real and virtual images is that if a screen is placed at the location of a real image, the image can be seen on the screen. Placing a screen at the location of a virtual image does not display the image because the light rays do not actually pass through the image location.

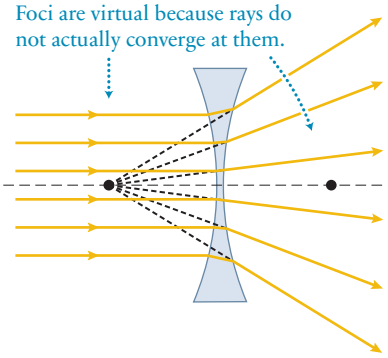
Figure 33.31*a* shows that, for this configuration of object and lens (object between focus and lens), the image is larger than the object and upright (unlike the image in Figure 33.28, which is inverted). A magnifying glass is designed to produce an enlarged, upright image of an object, which means that magnifying glasses are made with converging lenses and are held close to the object of interest (so that the object is between the lens and the focus).

33.11 As the object in Figure 33.31 is moved closer to the lens, does the size of the image increase, decrease, or stay the same?

Just as with electric circuits, it is convenient to use a simplified notation for ray diagrams. Figure 33.31*b* shows such a simplified version of the ray diagram of Figure 33.31*a*. Note that objects and real images are denoted by solid arrows and virtual images are denoted by dashed arrows.

Lenses can also be made with concave spherical surfaces rather than convex ones, where a *concave surface* is one that curves like the inside of a sphere. Such a lens is called a *diverging lens*, and **Figure 33.32** shows why: A series of parallel rays entering the lens are no longer parallel when they emerge. If we follow the path of the emerging rays back to the left side of the lens, we see that the

Figure 33.32 A diverging lens.



CONCEPTS

PROCEDURE: Principal rays for lenses

The propagations of principal rays for converging and diverging lenses are very similar. The description below holds for rays that travel from left to right.

Converging lens

1. A ray that travels parallel to the lens axis before entering the lens goes through the right focus after exiting the lens.
2. A ray that passes through the center of the lens continues undeflected.
3. A ray that passes through the left focus travels parallel to the lens axis after exiting the lens. If the object is between the focus and the lens, this ray doesn't pass through the focus but lies on the line from the focus to the point where the ray originates.

Diverging lens

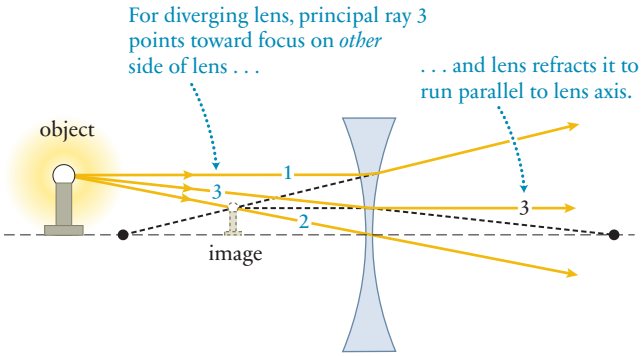
1. A ray that travels parallel to the lens axis before entering the lens continues along the line from the left focus to the point where the ray enters the lens.
2. A ray that passes through the center of the lens continues undeflected.
3. A ray that travels toward the right focus travels parallel to the lens axis after exiting the lens.

diverging rays appear to all come from the same location on the left side. This location corresponds to the focus of a converging lens, but in a diverging lens it is a *virtual focus* rather than a real focus because the rays never actually travel through this location. (Just as for converging lenses, there is an equivalent focus on the other side of the lens.)

Figure 33.33 shows a ray diagram for a diverging lens. The same principal rays are drawn, but now ray 3 does not pass through the focus on the same side of the lens as the object. Instead, for diverging lenses ray 3 is drawn on the line that runs from the point where the ray originates to the focus on the other side of the lens; once refracted by the lens, this ray travels parallel to the lens axis. The general procedure for drawing ray diagrams for diverging lenses is still the same as the one for converging lenses (see the Procedure box "Simplified ray diagrams for lenses" on page 1175), but the drawing of principal rays is a little bit different (see the Procedure box "Principal rays for lenses" above).

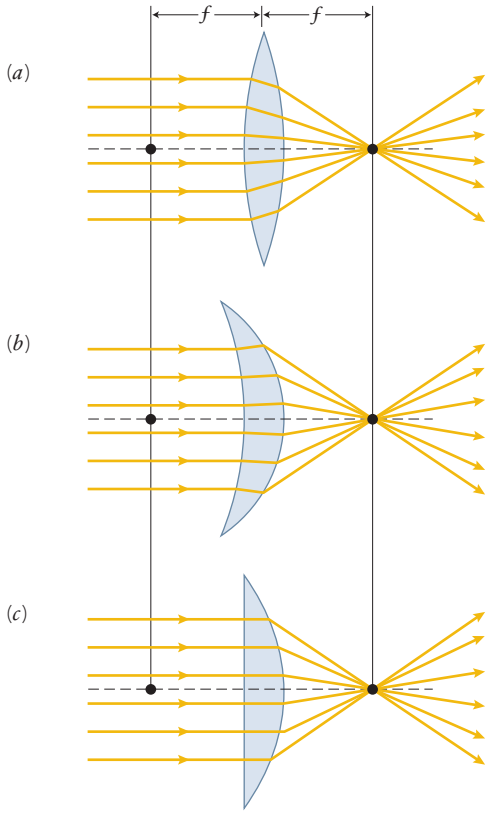
The lenses we have considered so far all have identical curved surfaces on each side. Many lenses have different surfaces, however. For example, it is possible to construct a converging lens with a certain focal length with two identical curved surfaces, two differently curved surfaces, or even a flat and a curved surface (**Figure 33.34**), as long as the lens is thicker at its center than at the edges. Regardless of the *radii of curvature* (that is, the radii of the spheres that best fit the surfaces), the lens has two foci, one on either side of the lens at the same distance f from it.

Figure 33.33 Ray diagram for an object outside the focus of a diverging lens.



33.12 Is the image in Figure 33.33 real or virtual?

Figure 33.34 These lenses are all converging because each is thicker at the center than at the edges.



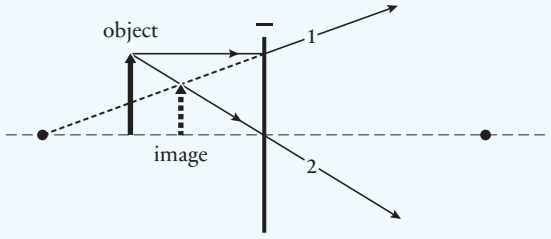
CONCEPTS

EXAMPLE 33.5 Demagnifying glass

Suppose the object in Figure 33.33 is placed between the focus and the lens. (*a*) Is the image real or virtual? (*b*) Is it larger than, smaller than, or the same size as the object?

❶ **GETTING STARTED** To sketch the situation (**Figure 33.35**), I represent the diverging lens as a vertical line with a minus sign above it, and draw the horizontal lens axis. I add the focal points at equal distances from the lens along its axis. Finally, I add a solid arrow, representing the object, between the left focal point and the lens.

Figure 33.35



❷ **DEVISE PLAN** To locate the image and determine its size and whether it is real or virtual, I can draw the principal rays.

❸ **EXECUTE PLAN** (*a*) I add to my diagram the principal rays coming from the tip of my object arrow. All I need is rays 1 and 2; I do not need ray 3 because the intersection of ray 2 and the dashed extension of ray 1 unambiguously determines the location of the tip of the image arrow. The dashed extension I had to draw for ray 1 tells me that the rays do not actually intersect at this location; they only appear to intersect here. Therefore the image is virtual. ✔

(*b*) My diagram tells me that the image is smaller than the object. ✔

❹ **EVALUATE RESULT** Because a diverging lens spreads rays out rather than bringing them together, it makes sense that a virtual image will form. I also know from experience that in contrast to a converging lens, which magnifies images, diverging lenses create smaller images, as I found.



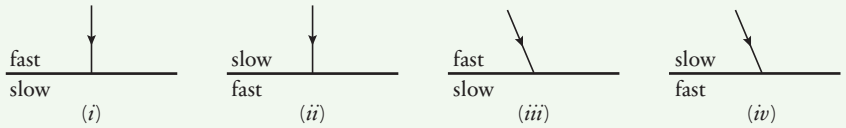
33.13 (*a*) Draw the third principal ray in Figure 33.35. Is there any position for the object in Figure 33.35 for which (*b*) the image is larger than the object and (*c*) the image is real?

CONCEPTS

Self-quiz

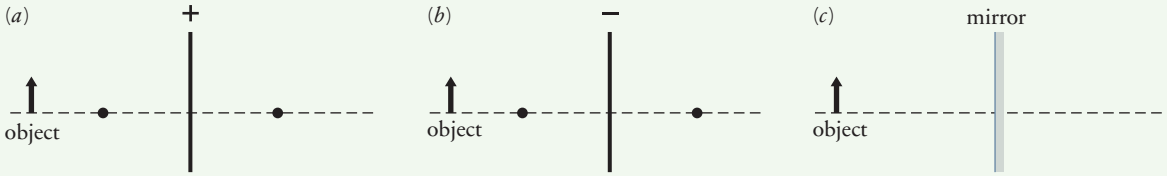
1. Why do you get a clear reflection from the surface of a lake on a calm day but little or no reflection from the surface on a windy day?
2. (*a*) As light travels from one medium into another, as shown in **Figure 33.36** ("fast" and "slow" refer to the wave speed in each medium), what happens to the wavelength of the light? (*b*) Draw the reflected and refracted rays at each surface.

Figure 33.36



3. What is the difference between a real image and a virtual image?
4. In each situation in **Figure 33.37**, draw the three rays emanating from the top of the object and reflecting or refracting from the optical element shown. Show the image, and state whether it is real or virtual.

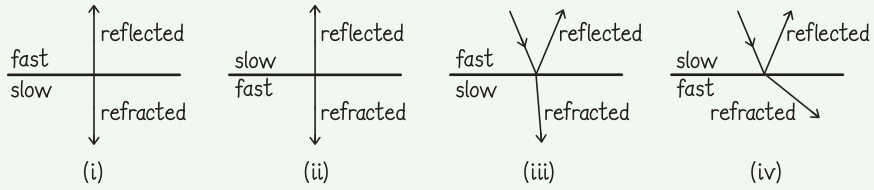
Figure 33.37



Answers

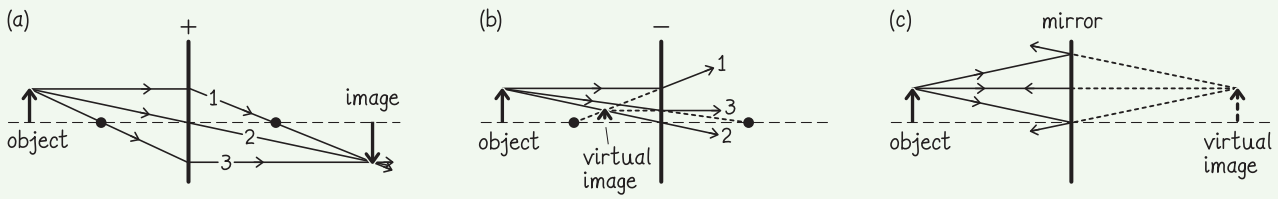
1. On a calm day, the lake surface is smooth, and specular reflection is like that of a mirror. On a windy day, the surface is rough, which makes the reflection diffuse and prevents the formation of an image.
2. (*a*) The wavelength decreases when the wave travels more slowly in the second medium (*i* and *iii*) and increases when the wave travels faster in the second medium (*ii* and *iv*). (*b*) See **Figure 33.38**.

Figure 33.38



3. Real image: All rays actually pass through the location of the image, and the image can be seen on a screen placed at the image location. Virtual image: All rays do not pass through the location of the image (only the extensions of the rays do), and the image cannot be seen on a screen placed at the image location.
4. See **Figure 33.39**. For the lenses, the three principal rays can be used to locate the image; for the mirror, any rays and the law of reflection can be used to locate the image. The images are (*a*) real, (*b*) virtual, (*c*) virtual.

Figure 33.39



33.5 Snel's law

In the first part of this chapter, we saw that light refracts when it travels from one medium into another because the speed of light depends on the medium. The speed of light in a medium is specified by the **index of refraction:**

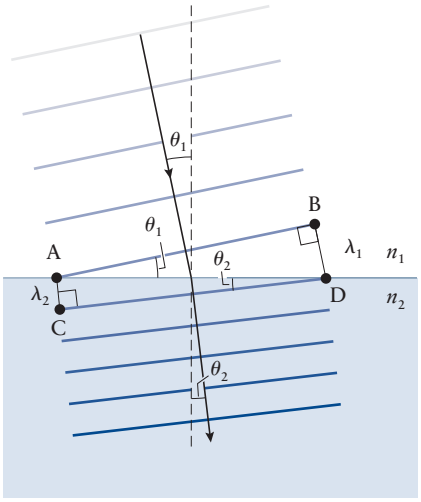
$$n \equiv \frac{c_0}{c}, \tag{33.1}$$

where c is the speed of light in the medium and c_0 is the speed of light in vacuum. (By definition, $n_{\text{vacuum}} = 1$; in air $n_{\text{air}} \approx 1$.) If a light wave of frequency f travels from one medium into another, the frequency doesn't change because the source determines the frequency (see also Checkpoint 16.10). The wavelength, however, does change; it is greater in the medium in which wave speed is greater.

The wavelength λ of the light is related to the wave speed and frequency, in the same manner that these quantities are related for harmonic waves (Eq. 16.10). In vacuum, for example,

$$\lambda = \frac{c_0}{f} \quad \text{(vacuum)}. \tag{33.2}$$

Figure 33.40 Relationship between angle of incidence θ_1 and angle of refraction θ_2.



In a medium in which a wave has speed c_1, the wavelength λ_1 is given by

$$\lambda_1 = \frac{c_1}{f} = \frac{c_0/n_1}{f} = \frac{1}{n_1}\lambda, \tag{33.3}$$

where λ is the wavelength of the wave in vacuum and n_1 is the index of refraction of the medium. Thus, the wavelength decreases as the index of refraction increases. As discussed in Section 33.3, the amount of refraction a light wave undergoes varies somewhat with wavelength (see Figure 33.22) because different wavelengths of light travel at different speeds. Therefore the index of refraction depends on the wavelength. **Table 33.1** lists the indices of refraction for some common transparent materials at a wavelength of 589 nm.

Let us now work out the quantitative relationship between the angle of incidence and the angle of refraction. **Figure 33.40** shows wavefronts and one ray for a beam of light incident on the interface between medium 1 and medium 2 at angle θ_1 from the normal. The angle of refraction is θ_2. Using right triangles ABD and ACD, we can express angles θ_1 and θ_2 in terms of the wavelengths λ_1 and λ_2:

$$\sin\theta_1 = \frac{\text{BD}}{\text{AD}} = \frac{\lambda_1}{\text{AD}} \tag{33.4}$$

and

$$\sin\theta_2 = \frac{\text{AC}}{\text{AD}} = \frac{\lambda_2}{\text{AD}}. \tag{33.5}$$

Table 33.1 Indices of refraction for common transparent materials

Material	n (for λ = 589 nm)
Air (at standard temperature and pressure)	1.00029
Liquid water	1.33
Sugar solution (30%)	1.38
Sugar solution (80%)	1.49
Microscope cover slip glass	1.52
Sodium chloride (table salt)	1.54
Flint glass	1.65
Diamond	2.42

QUANTITATIVE TOOLS

Combining these equations to eliminate AD and substituting $\lambda_1 = \lambda/n_1$ (Eq. 33.3), we get

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{\lambda_1}{\lambda_2} = \frac{\lambda/n_1}{\lambda/n_2} = \frac{n_2}{n_1}, \tag{33.6}$$

which can be written as

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{33.7}$$

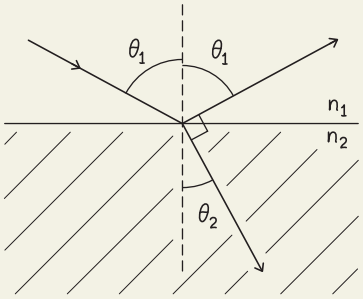
This relationship between the indices of refraction and the angles of incidence and refraction is called **Snel's law,** after the Dutch astronomer and mathematician Willebrord Snel van Royen (1580–1626). (In English, his name is often spelled as "Snell" instead.)

EXAMPLE 33.6 Bending 90°

A ray traveling through a medium for which the index of refraction is n_1 is incident on a medium for which the index of refraction is n_2. At what angle of incidence θ_1, expressed in terms of n_1 and n_2, must the ray strike the interface between the two media for the reflected and transmitted rays to be at right angles to each other?

❶ **GETTING STARTED** This problem involves both reflection and refraction at an interface between two media. To visualize the problem, I draw the incident, reflected, and refracted rays and indicate that the reflected and refracted rays are 90° apart (**Figure 33.41**).

Figure 33.41



❷ **DEVISE PLAN** Snel's law (Eq. 33.7), the law of reflection, and the indices of refraction determine the paths taken by the reflected and refracted rays. Therefore I need to use those relationships to obtain an expression that tells me the value of θ_1 that produces reflected and refracted rays oriented 90° to each other. To obtain θ_1 in terms of n_1 and n_2, I need to eliminate θ_2 from Eq. 33.7. To do so, I use the fact that the angles on the right side of the normal to the interface must add to 180°. Thus, with reflected and refracted rays forming a 90° angle, I can say $180° = \theta_1 + 90° + \theta_2$. Solving this expression for θ_2 gives $\theta_2 = 90° - \theta_1$, which I can substitute into Eq. 33.7.

❸ **EXECUTE PLAN** Substituting $\theta_2 = 90° - \theta_1$ into Eq. 33.7, I get

$$n_1 \sin\theta_1 = n_2 \sin(90° - \theta_1) = n_2 \cos\theta_1,$$

and isolating the terms that contain θ_1 gives

$$\frac{\sin\theta_1}{\cos\theta_1} = \tan\theta_1 = \frac{n_2}{n_1}$$

$$\theta_1 = \tan^{-1}\left(\frac{n_2}{n_1}\right). \checkmark$$

❹ **EVALUATE RESULT** My result says that θ_1 increases as n_2 increases. This makes sense because as n_2 increases, the refracted ray bends more, meaning that θ_2 becomes smaller. To keep the reflected and refracted rays perpendicular to each other, the angle of reflection must increase, and so θ_1 must also increase.

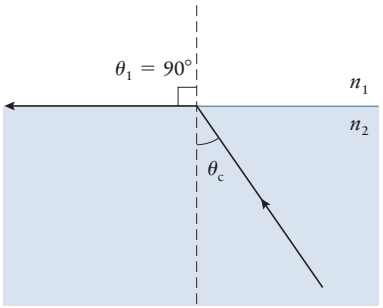
Earlier in this chapter, we found that for rays traveling from a denser medium to a less dense medium, we can define a critical angle of incidence θ_c such that the angle of refraction θ_1 is equal to 90° (**Figure 33.42**); beyond this critical angle θ_c, total internal reflection occurs. We can calculate the critical angle θ_c for an interface between two media with indices of refraction n_1 and n_2 ($n_2 > n_1$) by applying Snel's law (Eq. 33.7) and setting $\theta_1 = 90°$:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{1}{\sin\theta_c} = \frac{n_2}{n_1}. \tag{33.8}$$

Solving for θ_c gives

$$\theta_c = \sin^{-1}\left(\frac{n_1}{n_2}\right). \tag{33.9}$$

Figure 33.42 Critical angle for a ray traveling from a denser medium (n_2) to a less dense medium ($n_1 < n_2$).

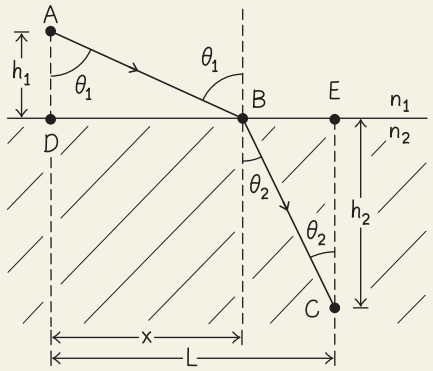


EXAMPLE 33.7 Fermat's principle

For a light ray that crosses the interface between medium 1 having index of refraction n_1 and medium 2 having index of refraction n_2, what relationship between θ_1 and θ_2 follows from Fermat's principle (page 1173)?

❶ **GETTING STARTED** I begin with a sketch of both media and a ray traveling from an arbitrary point A in the n_1 medium to an arbitrary point C in the n_2 medium (**Figure 33.43**). Fermat's principle states that the path the ray takes from A to C is the path for which the time interval needed for the motion is a minimum. Therefore this ray must cross the interface at a point B that makes the time interval a minimum.

Figure 33.43



An alternative way to express this problem is: Given the locations of A and C, where must B lie so as to minimize the time interval needed to travel from A to C?

❷ **DEVISE PLAN** I add two more location labels to my drawing: D directly below A and lying on the interface, and E directly above C and lying on the interface. Doing so gives me two right-angle triangles that permit me to express the angles in terms of the distance traveled. I write h_1 for the distance AD and h_2 for the distance EC. I can think of the distance from D to B as unknown—I'll call it x—and the distance from D to E, which I'll call L, is fixed by the locations of A and C. My goal is to determine the value of x for which the travel time from A to C is minimized. Once I obtain x, I hope to obtain a relationship between θ_1 and θ_2.

❸ **EXECUTE PLAN** I begin by expressing the time interval Δt_{AC} the ray needs to travel from A to C in terms of the distances shown in Figure 33.43 and the speed of light in the two media:

$$\Delta t_{AC} = \Delta t_{AB} + \Delta t_{BC} = \frac{AB}{c_1} + \frac{BC}{c_2} = \frac{AB}{c_0/n_1} + \frac{BC}{c_0/n_2}. \quad (1)$$

Next I express AB and BC in terms of h_1, h_2, L, and x:

$$AB = \sqrt{h_1^2 + x^2}$$

$$BC = \sqrt{h_2^2 + (L-x)^2}.$$

Substituting these two expressions into Eq. 1 gives me

$$\Delta t_{AC} = \frac{\sqrt{h_1^2 + x^2}}{c_0/n_1} + \frac{\sqrt{h_2^2 + (L-x)^2}}{c_0/n_2}.$$

Except for x, all quantities in this expression are constants.

The path for which the time interval Δt_{AC} is a minimum—as it must be from Fermat's principle—is the path for which the derivative of Δt_{AC} with respect to x is zero:

$$\frac{d}{dx}(\Delta t_{AC}) = \frac{xn_1}{c_0\sqrt{h_1^2 + x^2}} - \frac{(L-x)n_2}{c_0\sqrt{h_2^2 + (L-x)^2}} = 0. \quad (2)$$

Solving this equation for x would tell me where the light ray crosses the interface, but I do not have values for L and x. However, the right triangles ADB and BEC in Figure 33.43 allow me to express these distances in terms of θ_1 and θ_2:

$$\sin\theta_1 = \frac{x}{\sqrt{h_1^2 + x^2}} \quad \text{and} \quad \sin\theta_2 = \frac{L-x}{\sqrt{h_2^2 + (L-x)^2}}. \quad (3)$$

I can now use these expressions to rewrite Eq. 2 in terms of θ_1 and θ_2. From Eq. 2 I obtain

$$\frac{xn_1}{c_0\sqrt{h_1^2 + x^2}} = \frac{(L-x)n_2}{c_0\sqrt{h_2^2 + (L-x)^2}}.$$

Canceling the c_0 factors that appear on both sides and substituting from Eq. 3, I get

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \checkmark$$

❹ **EVALUATE RESULT** My result is identical to Snel's law—which I derived by considering the effect of changing speed on the propagation of wavefronts. So Fermat's principle yields the same result as Snel's law, which I know to be correct.

Fermat's principle applies to all of ray optics, not only to refraction. As discussed in the box "Fermat's principle" on page 1173, the law of reflection also follows from this principle.

33.14 We found in Example 33.3 that a light ray is refracted twice when it passes completely through a slab of transparent material (see Figure 33.18). The result of these two refractions is that the exiting ray is shifted sideways relative to the entering ray. Let the slab be in air with an index of refraction $n_1 = 1$. (*a*) Derive an expression for the distance (perpendicular to the ray) over which the ray is shifted sideways for an angle of incidence θ_1, slab thickness d, and slab index of refraction n_2. (*b*) Calculate the value of the shift for $\theta_1 = 30°$, $n_2 = 1.5$, and $d = 0.010$ m.

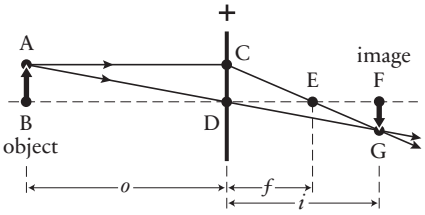
QUANTITATIVE TOOLS

33.6 Thin lenses and optical instruments

In the first part of this chapter, we found that converging lenses form images of objects by focusing the light rays emanating from those objects. Let us now work out quantitatively the location and size of such images. We shall restrict our discussion to lenses that are thin enough that we can ignore the type of effects shown in Figure 33.29. Such lenses are called *thin lenses*.

A simplified ray diagram of the image formed by a converging lens is shown in **Figure 33.44**. The focal length f of the lens is DE, the distance o from the lens to the object (also called the *object distance*) is BD, and the distance i from the lens to the image (also called the *image distance*) is DF. The height of the object is AB, and the height of the image is FG. Let us denote the height of the object by h_o and the height of the image by h_i. We choose the values of h_o and h_i to be positive for upright objects and images and negative for inverted objects and images. We want to obtain a relationship between h_i and h_o, which will tell us how large the image is relative to the object. We also want a relationship among f, i, and o, which will tell us how the positions of the object and the image are related.

Figure 33.44 Simplified ray diagram for the formation of an image by a converging lens.



We begin by noting that triangles ABD and DFG are similar, which means

$$\frac{\text{AB}}{\text{DB}} = \frac{\text{FG}}{\text{DF}}. \tag{33.10}$$

Because the image is inverted, $h_i = -\text{FG}$, we can rewrite Eq. 33.10 as

$$\frac{h_o}{o} = \frac{-h_i}{i}. \tag{33.11}$$

Rearranging this expression gives

$$-\frac{h_o}{h_i} = \frac{o}{i}. \tag{33.12}$$

In this case, the absolute value of the ratio of the object height to the image height equals the ratio of the object distance to the image distance.

Triangles CDE and EFG are also similar, which means

$$\frac{\text{DE}}{\text{CD}} = \frac{\text{EF}}{\text{FG}}, \tag{33.13}$$

which can be written as

$$\frac{f}{h_o} = \frac{i-f}{-h_i}. \tag{33.14}$$

Using Eq. 33.12 to rewrite Eq. 33.14 in terms of f, o, and i gives us

$$\frac{f}{o} = \frac{i-f}{i} = 1 - \frac{f}{i}. \tag{33.15}$$

Dividing by f and rearranging terms yield

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i}. \tag{33.16}$$

This result is known as the **lens equation.**

It can be shown that Eq. 33.16 is generally true for either real or virtual images formed by either converging or diverging lenses, as long as we choose the signs of f, i, and o properly. For a converging lens, f is positive and o is positive if the object is in front of the lens. (This is always true for a single lens and for the first lens in a lens combination. For situations involving multiple lenses, however, it is possible that the object imaged by a secondary lens is on the opposite side of the lens from the side where the rays enter it—that is, the object is "behind the lens." In that case o is negative.) If the image is on the same side of the lens as the emerging light, the image is real and i is positive; if the image is on the opposite side of the lens from the emerging light, the image is virtual and i is negative.

For a diverging lens (a lens with concave surfaces), the focal length f is negative because the focus is virtual rather than real—that is, parallel rays appear to come from the same side of the lens where the light source is rather than converging on the other side of the lens. The same sign convention applies to o as for converging lenses. A single diverging lens always produces a virtual image (see Example 33.5), so i is always negative for such lenses.

The sign conventions for f, i, and o are similar for images formed by spherical mirrors, which are discussed in the next section. **Table 33.2** summarizes these sign conventions.

The **magnification** of the image is defined as the ratio of the signed image height to the object height. Using Eq. 33.12, we get

$$M \equiv \frac{h_i}{h_o} = -\frac{i}{o}. \tag{33.17}$$

We define M this way so that the magnification of upright images is positive and that of inverted images is negative. Examining Figures 33.30, 33.31, and 33.33, we can see that for a single lens, when the image distance and object distance are both positive, as in Figure 33.30, we obtain an inverted image, whereas when the image distance is negative, as in Figures 33.31 and 33.33, we obtain an upright image.



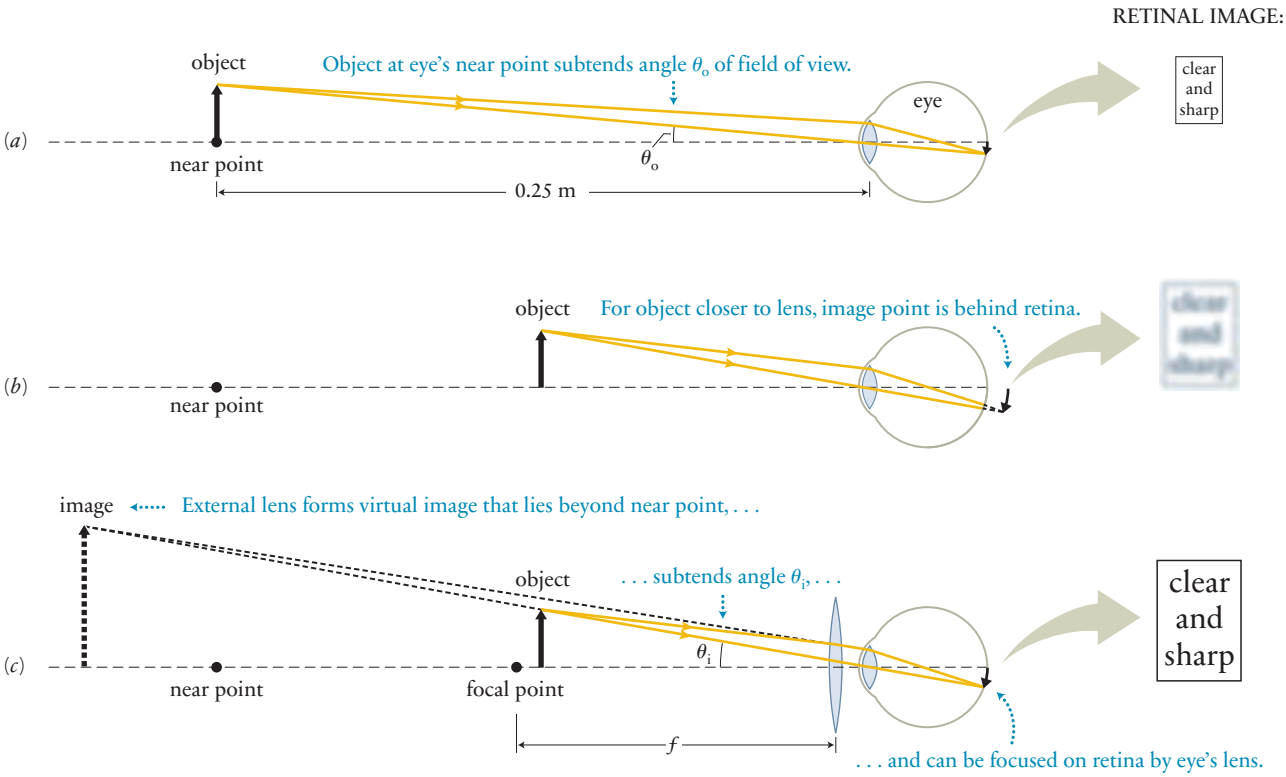
33.15 If the diverging lens in Figure 33.33 has a focal length of 80 mm and the object is located 100 mm from the lens, (*a*) what is the image distance and (*b*) how tall is the image relative to the object?

Table 33.2 Sign conventions for f, i, and o (positive = real; negative = virtual)

Sign	Lens	Mirror
$f > 0$	converging lens	converging mirror
$f < 0$	diverging lens	diverging mirror
$o > 0$	object in front[b] of lens	object in front of mirror
$o < 0$[a]	object behind lens	object behind mirror
$i > 0$	image behind lens	image in front of mirror
$i < 0$	image in front of lens	image behind mirror
$h_i > 0$	image upright	image upright
$h_i < 0$	image inverted	image inverted
$\|M\| > 1$	image larger than object	image larger than object
$\|M\| < 1$	image smaller than object	image smaller than object

[a]Encountered only with lens or mirror combinations.
[b]For both lenses and mirrors, "in front" means on the side where the rays originate; "behind" refers to the opposite side.

Figure 33.45 An eye cannot focus on an object that is closer than its near point (which represents the limit of the biological lens's ability to change curvature). However, an external converging lens (such as a magnifying lens) makes it possible to see objects that are closer than the near point. It also enlarges them.



The human eye focuses incoming light rays, forming an image on the retina of the eye (**Figure 33.45*a***). One part of the eye is its lens, but unlike the lenses we have examined so far, the focal length of the eye's lens is variable, which allows us to see objects clearly over a wide range of distances. When the muscle around the lens is fully relaxed, the lens flattens out and the retina lies in the focal plane of the lens. Thus, light rays from distant objects focus onto the retina.

To the unaided eye, the largest (and thus most detailed) image of an object is observed when we bring the object as close as possible to the eye. However, there is a limit to how much the eye's lens can adjust. The *near point* is the closest object distance at which the eye can focus on the object comfortably. Typically, for an adult, the near point is about 0.25 m from the eye. An object positioned at the near point appears clear and sharp to the observer, as shown in Figure 33.45*a*. With age the distance between the near point and the eye tends to increase, and when an object is brought closer than the near point, the plane where the image is formed lies behind the retina and the image "seen" by the retina is blurry (Figure 33.45*b*). This situation can be corrected by an external lens that works in combination with your eye's lens to focus the image on the retina (Figure 33.45*c*).

An external converging lens properly placed between the object and the eye, as in Figure 33.45*c*, magnifies the object. To maximize the size of the image, the object is held near the focus of the external lens, and the lens is held as close as possible to the eye. The image formed by the external lens then serves as the object for the eye's lens. The image formed by the external lens is virtual and subtends an angle θ_i that is greater than the angle θ_o subtended by the object in Figure 33.45*a*, permitting the viewer to see finer

details. The image is also outside the near point; if the object is placed exactly at the focus of the external lens, the image is at infinity and can be viewed comfortably. We can define the *angular magnification* produced by the lens as

$$M_\theta \equiv \left|\frac{\theta_i}{\theta_o}\right|. \tag{33.18}$$

For small angles and an object placed close to the focus of the external lens, as in Figure 33.45*c*, the angle θ_i subtended by the image can be expressed in terms of the object height h_o and the focal length f of the lens:

$$\theta_i \approx \tan\theta_i \approx \frac{h_o}{f} \quad \text{(object close to focus, small } \theta_i\text{)}. \tag{33.19}$$

For small angles and an object placed at the eye's near point, as in Figure 33.45*a*, the angle subtended by the object is approximately

$$\theta_o \approx \tan\theta_o = \frac{h_o}{0.25\ \text{m}} \quad \text{(object close to near point, small } \theta_o\text{)}. \tag{33.20}$$

Substituting Eqs. 33.19 and 33.20 into Eq. 33.18 gives an angular magnification of

$$M_\theta \approx \frac{0.25\ \text{m}}{f}. \tag{33.21}$$

This expression gives what is called either the *small-angle approximation* or the *paraxial approximation* to the angular magnification because it is obtained with the small-angle approximations of Eqs. 33.19 and 33.20. These approximations are good to within 1% for angles of 10° or less.

Lenses placed near the eye (in the form of eyeglasses) are used to correct vision for far-sighted or near-sighted eyes. The strength of eyeglass lenses (and of magnifying lenses, too) is commonly symbolized by d and measured in *diopters*:

$$d \equiv \frac{1\ \text{m}}{f}. \tag{33.22}$$

The *lens strength* d, like the lens focal length f, is positive for converging lenses and negative for diverging lenses. For example, a +4-diopter lens is a converging lens with a focal length of 0.25 m. Diverging lenses are typically used to correct nearsightedness, with lens strengths ranging from −0.5 to −4 diopters.

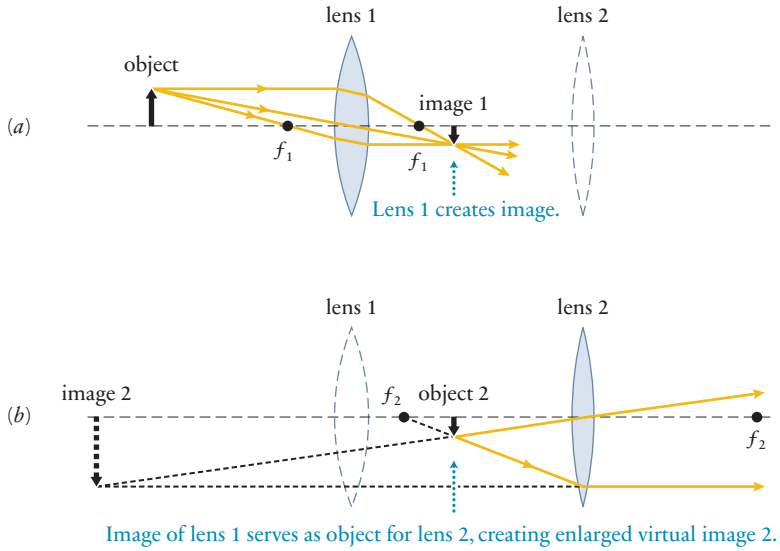


33.16 A single-lens magnifying glass used to examine photographic slides produces eightfold angular magnification. (*a*) What is the lens strength in diopters? (*b*) What is the focal length of the lens?

Many optical instruments combine two or more lenses to increase magnification. To trace rays through a combination of lenses, use the following procedure: The image formed by the first lens serves as the object for the second lens, the image formed by the second lens serves as the object for the third lens, and so on. **Figure 33.46** shows a ray diagram constructed in two steps for a combination of two lenses. Figure 33.46*a* shows the object, image, and rays for lens 1, and Figure 33.46*b* shows these elements for lens 2. Note that the rays from object 2 (which is the image formed by lens 1) are *not* the continuation of those used to locate image 1.

QUANTITATIVE TOOLS

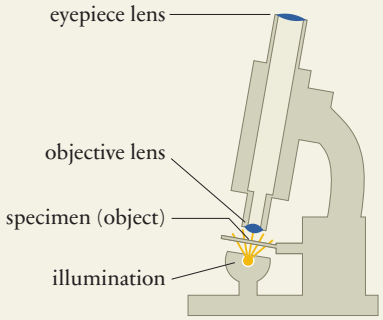
Figure 33.46 Two-step process for tracing rays through a combination of two lenses. When lenses are combined, the image of each lens serves as an object for the next lens.



EXAMPLE 33.8 Compound microscope

A compound microscope consists of two converging lenses, the *objective lens* and the *eyepiece lens*, positioned on a common optical axis (**Figure 33.47**). The objective lens is positioned to form a real, highly magnified image 1 of the sample being examined, and the eyepiece lens is positioned to form a virtual, further magnified image 2 of image 1. It is image 2 that the user sees. A knob on the microscope allows the user to move the objective lens upward and downward to change both the sample-objective lens distance and the distance between the two lenses. (*a*) How must the sample and the two lenses be positioned relative to one another so that the user sees a highly magnified, virtual image of the sample? (*b*) What is the overall magnification produced by the microscope?

Figure 33.47 Example 33.8.



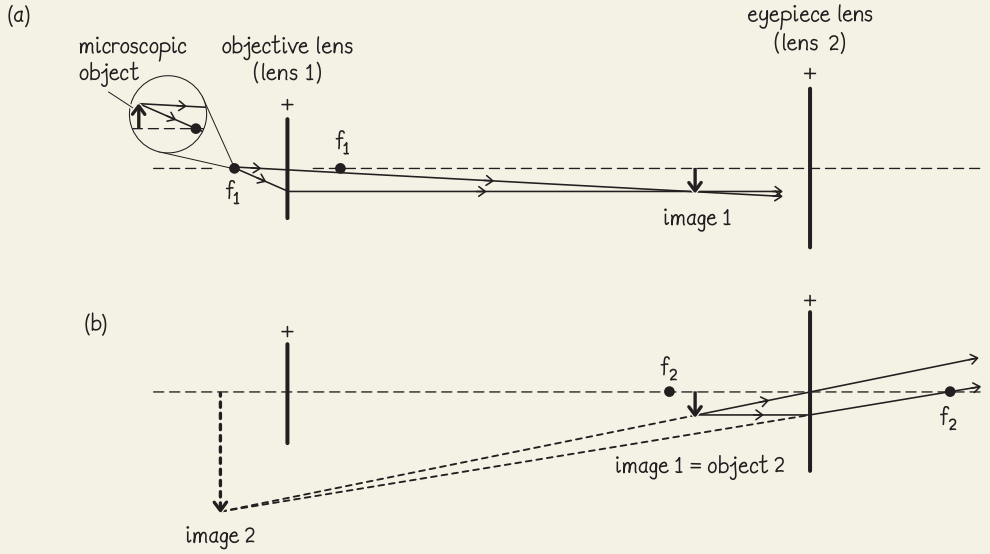
❶ **GETTING STARTED** I begin by examining Figure 33.46, which shows how, in a combination of two lenses 1 and 2, the image formed by lens 1 serves as the object for lens 2. The objective lens in a compound microscope corresponds to lens 1 in Figure 33.46, and the eyepiece lens corresponds to lens 2. Thus to keep things simple I refer to the objective lens as 1 and the eyepiece lens as 2.

❷ **DEVISE PLAN** To determine the relative positioning of the two lenses relative to each other, I must examine ray diagrams for various lens-sample distances and determine for which arrangement I get the greatest magnification. To determine the magnification M_1 of image 1, I can use the lens equation (Eq. 33.16) together with the relationship among magnification, image distance, and object distance (Eq. 33.17). The focal length of the lenses is fixed by their construction, and in operating a microscope, the observer can adjust the distance between the sample and the objective lens, so I can express this magnification in terms of f_1 and o_1. To determine the magnification of image 2, I recognize that lens 2 is used as a magnifying glass, so I can use Eq. 33.21, which gives the angular magnification $M_{\theta 2}$ produced by a simple magnifier. The overall magnification produced by the microscope is the product $M_1 M_{\theta 2}$.

❸ **EXECUTE PLAN** (*a*) In Figure 33.46, lens 1 produces an image that is smaller than the object. I am told that the image formed by lens 1 in a microscope is larger than the sample, and so I must choose a different sample position, one that yields an image 1 larger than the sample. I am also told that this image is real. Placing the sample just outside the focal point of lens 1 gives me an image 1 that is larger than the sample. My choices are to increase or decrease the sample-lens 1 distance. Drawing a ray diagram for each possibility, I see that moving the sample farther and farther from lens 1 makes the image smaller and smaller. Therefore I should position the sample closer to lens 1 than in Figure 33.46*a*. Should I choose a position inside or outside the lens focus? I know

(Continued)

Figure 33.48



from the problem statement that this image is real, and I know from Figure 33.31 that an object inside the focus of a converging lens produces a virtual image. Thus my best choice is to adjust the sample-lens 1 distance so that the sample is just outside the lens focus (**Figure 33.48a**).

I am told that image 2 is virtual and larger than image 1. I know from Figure 33.31 that a converging lens produces a virtual, magnified image when the object is inside the lens focus. I again draw ray diagrams for various positions inside the focus and see that the greatest magnification is obtained when I adjust the distance from lens 1 to lens 2 to make image 1 fall just inside the focal point of lens 2, as shown in Figure 33.48*b*. ✔

(*b*) To determine M_1, I use the lens equation to write i_1 in terms of f_1 and o_1:

$$i_1 = \frac{1}{\left(\frac{1}{f_1} - \frac{1}{o_1}\right)}.$$

I substitute this expression into Eq. 33.17:

$$M_1 = -\frac{i_1}{o_1} = -\frac{1}{o_1} \times \frac{1}{\left(\frac{1}{f_1} - \frac{1}{o_1}\right)} = -\frac{1}{\frac{o_1}{f_1} - 1},$$

which tells me that the magnification M_1 produced by lens 1 is determined by the ratio o_1/f_1. Because I have made o_1 slightly larger than f_1 in order to produce a real image 1, the denominator is positive and therefore M_1 is negative.

The angular magnification produced by lens 2 is approximately

$$M_{\theta 2} = \frac{0.25\ \text{m}}{f_2}.$$

The overall magnification produced by the microscope is thus

$$M = M_1 M_{\theta 2} = \frac{-0.25\ \text{m}}{f_2\left(\frac{o_1}{f_1} - 1\right)}. \checkmark$$

4 EVALUATE RESULT Figure 33.48*a* indicates that image 1 is inverted, making M_1 negative and giving me confidence in my expression for M_1. Figure 33.48*b* tells me that image 2 is upright relative to its object, and so $M_{\theta 2}$ is positive, which agrees with my result. Because image 2 is inverted relative to the sample, the overall magnification is negative, as my result shows.

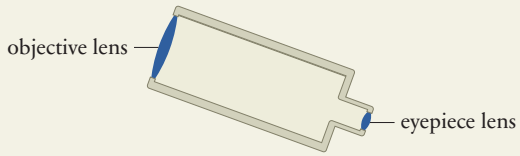
QUANTITATIVE TOOLS

33.17 (*a*) Consider replacing the objective lens in Fig. 33.48*a* with one that has a greater focal length, and moving the sample in order to keep it just outside the focal point of the lens. Does the image formed by the objective lens move closer to the objective lens, stay in the same place, or move farther from the objective lens? (*b*) In practice it is desirable for a microscope to be fairly compact. To keep the microscope compact, should the focal length of the objective lens be chosen to be short or long, or does it matter?

EXAMPLE 33.9 Refracting telescope

A refracting telescope, like a compound microscope, contains two converging lenses, the objective lens and the eyepiece lens, positioned on a common optical axis (**Figure 33.49**). However, a telescope is designed to view large, very distant objects, whereas a microscope is used to view very small objects that are placed very close to the objective lens. Consequently, the arrangement of lenses in a telescope is different from the arrangement in a microscope. The telescope's objective lens is positioned to form a real image of very distant objects, and the eyepiece lens is positioned to form a virtual image of the image produced by the objective lens, to be viewed by an observer. (*a*) How should the lenses be arranged to accomplish this? (*b*) What is the overall magnification produced by the telescope?

Figure 33.49 Example 33.9.



❶ **GETTING STARTED** I begin by examining Figure 33.46 and then construct a similar ray diagram with the object at a very great distance from the lenses. As in Example 33.8, I use lens 1 to refer to the objective lens and lens 2 to refer to the eyepiece lens. Because the object is very far away, light rays from it enter lens 1 as parallel rays. These rays form an image 1 in the focal plane of lens 1 (**Figure 33.50*a***). Because I know the location of image 1, I need to draw only one principal ray. As in the microscope, lens 2 is used as a simple magnifier to view image 1.

❷ **DEVISE PLAN** Because the original object is very distant and the final image is viewed by the observer's eye, I can calculate the angular magnification of this image. Although I could calculate the angular magnification produced by each lens and multiply them together, in this case it is simpler to determine the overall angular magnification because the angles θ_o and θ_i the object and image 2 subtend at the observer's eye are both very small. I can determine the overall angular magnification by taking the ratio of these angles while using the small-angle approximation.

❸ **EXECUTE PLAN** (*a*) In order for lens 2 to produce a magnified, virtual image of image 1, image 1 should be positioned just inside the focal plane of lens 2. If lens 2 is placed such that the image is at the focal plane of lens 2 (Figure 33.50*b*), lens 2 forms an infinitely distant, virtual image that can be viewed comfortably by the observer's relaxed eye. As my diagram shows, the lenses are then arranged such that their foci coincide. ✔

(*b*) **Figure 33.51*a*** shows the ray that passes through the foci of the lenses, labeled with the angles θ_o (subtended by the object) and θ_i (subtended by the image). Figure 33.51*b* shows the triangles I use to relate each of these angles to the height h_i of image 1 and the focal lengths of the lenses. The angular magnification is the ratio θ_o/θ_i. I can approximate

Figure 33.51

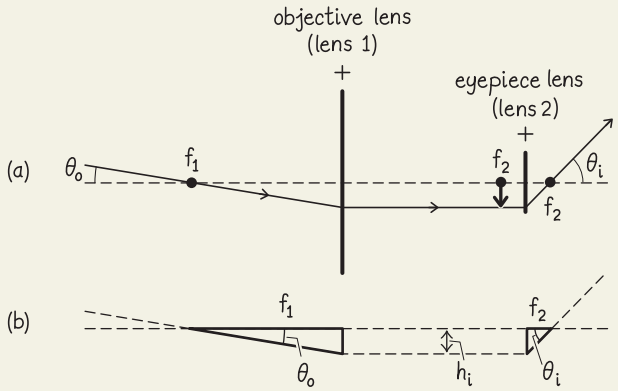
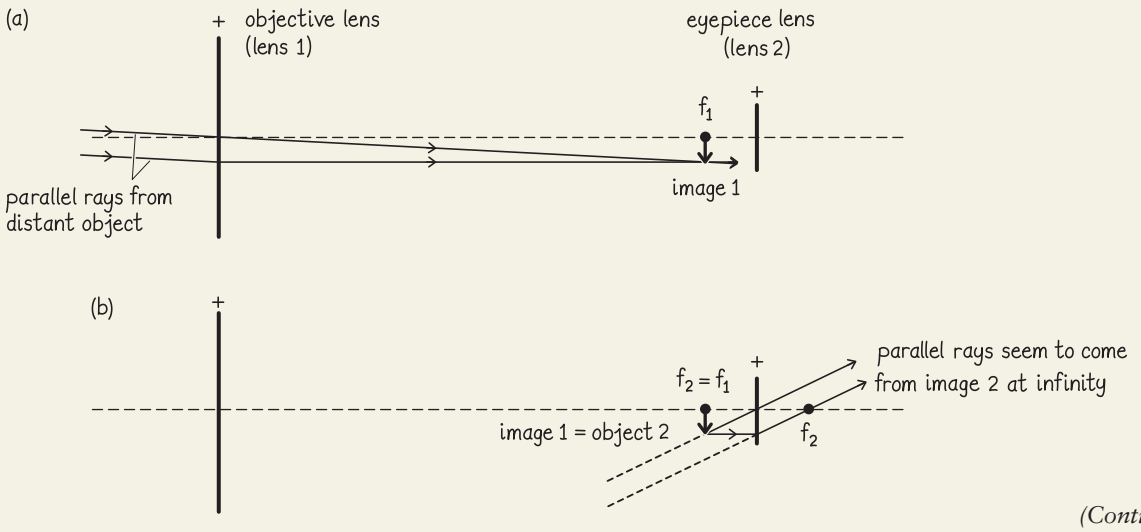


Figure 33.50



(Continued)

these angles by their tangents; substitute for the tangents of the angles in terms of h_i, f_1, and f_2; and simplify:

$$M_\theta = \left|\frac{\theta_i}{\theta_o}\right| \approx \left|\frac{\tan\theta_i}{\tan\theta_o}\right| = \frac{|h_i/f_2|}{|h_i/f_1|} = \left|\frac{f_1}{f_2}\right|. \checkmark$$

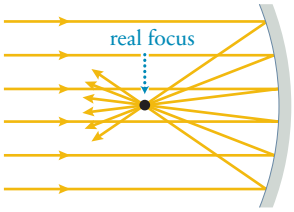
❹ **EVALUATE RESULT** Figure 33.50 indicates that image 1 is inverted and that image 2 is upright relative to image 1, which means that image 2 is inverted relative to the distant object. This makes sense because the incoming ray in Figure 33.51 is angled downward but the outgoing ray is angled upward.

33.18 A telescope with a magnification of 22× has an eyepiece lens for which the focal length f_2 is 40.0 mm. (*a*) What is the focal length f_1 of the objective lens? (*b*) What is the length of the telescope?

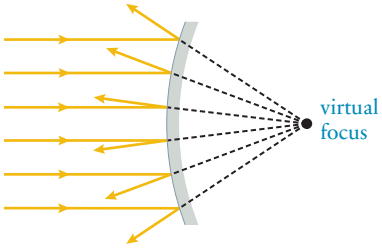
33.7 Spherical mirrors

Figure 33.52 Spherical mirrors focus parallel rays just like lenses do.

(*a*) Concave spherical mirror



(*b*) Convex spherical mirror



Just like lenses, spherical mirrors focus parallel rays (**Figure 33.52**). A concave mirror focuses rays to a point in front of the mirror, corresponding to a real focus; a diverging mirror makes rays diverge so that they appear to come from a point behind the mirror, corresponding to a virtual focus. Thus, just as with lenses, we can have both converging and diverging spherical mirrors. Unlike lenses, however, spherical mirrors have only a single focus.

To obtain the location of the focus of a converging mirror, we examine the reflection of the two rays shown in **Figure 33.53*a***. Ray 1 comes in parallel to the axis of the mirror, striking the mirror at A, and ray 2 comes in along the axis of the mirror. The focus of the mirror is at D, where the two reflected rays cross, and the center of the sphere on which the surface of the mirror lies is at C. Line CA is therefore a radius of the sphere and perpendicular to the mirror surface. We denote the length of the radius by R. This distance is called the *radius of curvature*.

Ray 2 strikes the mirror at normal incidence and so is reflected back along the axis. Ray 1 is reflected through the focus at D. Consequently, the distances CD and AD in Figure 33.53*a* are equal. Dividing triangle ACD into two congruent right triangles by drawing a line perpendicular to the base from D, we see that $CD = (R/2)/\cos\theta_i$. For small θ, $\cos\theta \approx 1$, and so $CD = R/2$ and $BD = R - CD = R/2$, independent of θ. Therefore, the focus, which is located at D because that is where the two reflected rays cross, lies halfway between the mirror and the center of the sphere, making the focal length

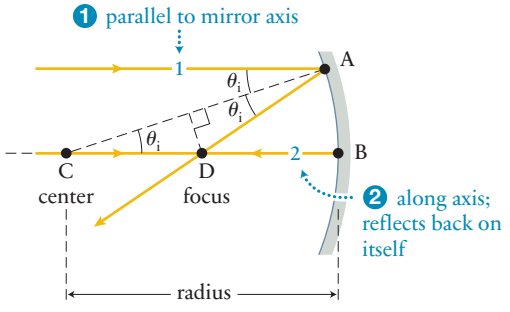
$$f = \frac{R}{2}. \tag{33.23}$$

As Figure 33.53*b* shows, the geometry and hence the position of the focus are exactly the same for diverging mirrors except that here the focus lies behind the mirror.

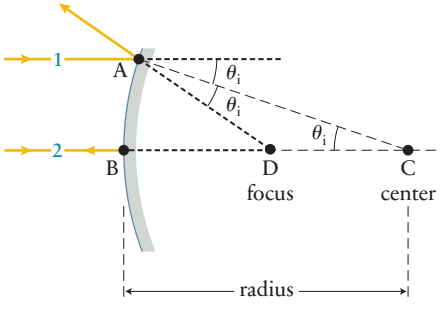
To determine the location of images formed by mirrors, we follow the same ray-tracing procedure we used for lenses. The three principal rays

Figure 33.53 The two principal rays used to determine the focus of a spherical mirror.

(*a*) Concave spherical mirror



(*b*) Convex spherical mirror



QUANTITATIVE TOOLS

PROCEDURE: Ray diagrams for spherical mirrors

Ray diagrams for spherical mirrors are very similar to those for lenses. This procedure is for rays traveling from the left to the right.

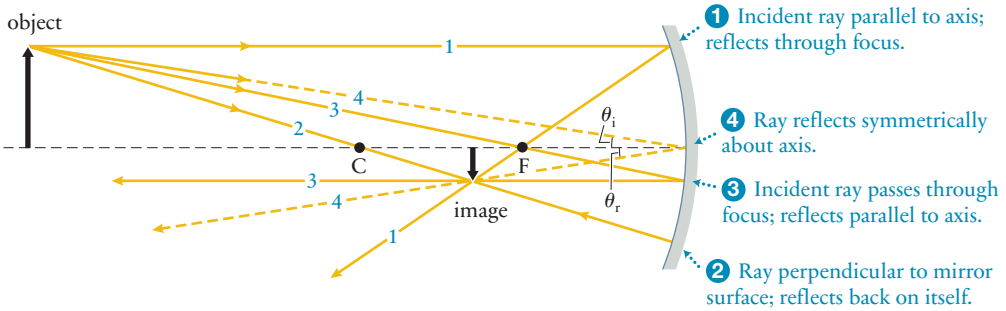
1. Draw a horizontal line representing the mirror axis. In the center of the diagram, draw a circular arc representing the mirror. A converging mirror curves toward the left; a diverging mirror curves toward the right.
2. Put a dot on the axis at the center of the circular arc and label it C. Add another dot on the axis, halfway between C and the mirror. This point is the focus. Label it f.
3. Represent the object by drawing an upward-pointing arrow from the axis at the appropriate relative distance to the left of the mirror. For example, if the distance from the object to a converging mirror is one-third the radius of curvature of the mirror, put the arrow a bit to the right of the focus. The top of the arrow should be at about half the height of the mirror.
4. From the top of the arrow representing the object draw two or three of the following three so-called *principal rays* listed in the Procedure box "Principal rays for spherical mirrors" on the next page.
5. The top of the image is at the point where the rays that are reflected by the mirror intersect. If the intersection is on the left side of the lens, the image is real; if it is on the right side, the image is virtual. Draw an arrow pointing from the axis to the intersection to represent the image (use a dashed arrow for a virtual image).

emanating from a given point on the object are analogous: ray 1 approaching the mirror parallel to the mirror axis, ray 2 passing through the center C of the mirror, and ray 3 passing through the focus on its way to the mirror. Now, however, "center" refers to *the center of the sphere on which the mirror surface lies* rather than the center of the lens. **Figure 33.54** shows a ray diagram for an image formed by a converging mirror. Ray 1 is reflected through the focus, ray 2 strikes the mirror at normal incidence and thus reflects back on itself, and ray 3 is reflected parallel to the mirror axis. As Figure 33.54 shows, there is a fourth ray that can easily be drawn: A ray that hits the mirror on the axis is reflected back symmetrically about the axis. The procedures for drawing ray diagrams and principal rays for spherical mirrors are given in the Procedure boxes on this page and the next.

Object distance, image distance, and focal length are measured from the surface of the mirror, and the relationship among o, i, and f is the same as that for lenses:

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i}. \tag{33.24}$$

Figure 33.54 Principal ray diagram for an object outside the focus of a concave spherical mirror. The image is real, inverted, and smaller than the object.



Just as for lenses, the focal length f for any spherical mirror is positive for a real focus and negative for a virtual focus, and o is positive when the object is in front of the mirror and negative when it is behind the mirror (that can happen only when the object is an image formed by another mirror or a lens). Similarly, i is positive for a real image and negative for a virtual image.

PROCEDURE: Principal rays for spherical mirrors

This description holds for rays that travel from left to right.

Converging mirror

1. A ray that travels parallel to the mirror axis before reaching the mirror goes through the focus after being reflected.
2. A ray that passes through the center of the sphere on which the mirror surface lies is reflected back onto itself. If the object is between the center and the mirror, this ray doesn't pass through the center but lies on the line from the center to the point at which the ray originates.
3. A ray that passes through the focus is reflected parallel to the axis. If the object is between the focus and the mirror, this ray doesn't pass through the focus but lies on the line from the focus to the point at which the ray originates.

Diverging mirror

1. A ray that travels parallel to the mirror axis before reaching the mirror is reflected along the line that goes through the focus and the point where the ray strikes the surface.
2. A ray that passes through the center of the sphere on which the mirror surface lies is reflected back onto itself.
3. A ray whose extension passes through the focus is reflected parallel to the axis.

For both converging and diverging mirrors, a ray that hits the mirror on the axis is reflected back symmetrically about the axis.

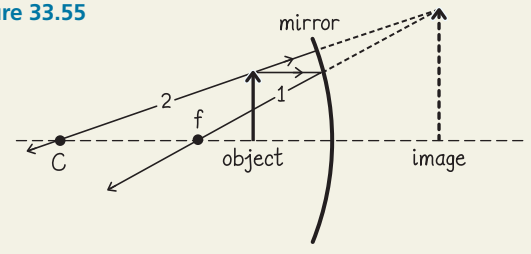
With mirrors, however, a real image is located on the same side of the mirror as the object and a virtual image is located on the opposite side—the opposite of what happens with lenses. These sign conventions are summarized in Table 33.2. Finally, the relationship between object and image distances and heights for lenses (Eq. 33.12) also applies to mirrors, so that equation can be used to determine the size of the images formed by mirrors.

EXAMPLE 33.10 Funny mirror

An object is placed 0.30 m in front of a converging mirror for which the radius of curvature is 1.0 m. (*a*) On which side of the mirror is the image? Is the image real or virtual? (*b*) If the object is 50 mm tall, what is the height of the image?

❶ **GETTING STARTED** To visualize the situation, I draw the mirror and its axis, and indicate its center of curvature C and its focal point *f* halfway between the mirror surface and C. I represent the object as a solid arrow (Figure 33.55).

Figure 33.55



❷ **DEVISE PLAN** To locate the image and identify whether it is real or virtual, I can draw the principal rays. I can use Eq. 33.12 to obtain the image height from the object height and the image and object distances. I can determine the image distance from the focal length and the object distance using Eq. 33.24.

❸ **EXECUTE PLAN** (*a*) As Figure 33.55 shows, I need to draw only two principal rays because the intersection of two rays unambiguously determines the position of the image. I draw ray 1 parallel to the mirror axis and reflecting through the focus. Because the object is between the mirror and the center of curvature, I draw ray 2 along the line defined by C and the tip of the object. Ray 2 does not pass through C until after it is reflected from the mirror, however, and I indicate this by adding an arrowhead pointing toward the mirror on the part of the ray to the right of the object and an arrowhead pointing away from the mirror on the part to the left of the object. My diagram shows that the rays do not actually meet but appear to come from a point behind the mirror. Therefore the image is behind the mirror and virtual. ✔

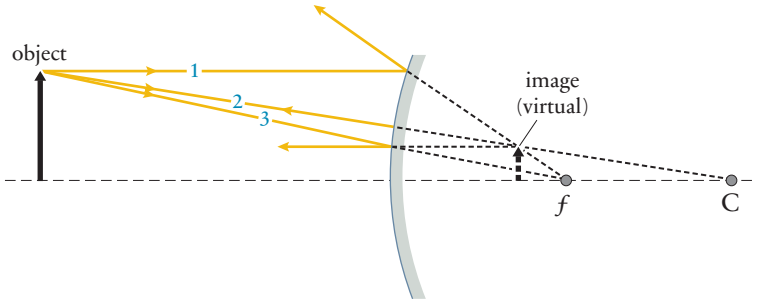
(*b*) I begin by determining the image distance *i*. The focal length of the mirror is half the radius of curvature, $f = 0.50$ m. Substituting this value and $o = 0.30$ m into Eq. 33.24 and solving for *i* give me $i = -0.75$ m. The negative sign indicates that the image is virtual and therefore behind the mirror; this is consistent with my result from part *a*. I then solve Eq. 33.12 for h_i (the signed image height) and substitute the values from this problem:

$$h_i = \frac{-i\,h_o}{o} = \frac{-(-0.75\text{ m})(0.050\text{ m})}{(0.30\text{ m})} = 0.13\text{ m}. \checkmark$$

❹ **EVALUATE RESULT** My ray diagram (Figure 33.55) indicates that the image is enlarged and upright, so h_i should be positive and greater than h_o. This agrees with my result.

QUANTITATIVE TOOLS

Figure 33.56 Ray diagram for an image formed by a convex spherical mirror (object distance greater than focal length). The image is virtual, upright, and smaller than the object.



Example 33.10 shows that placing an object inside the focus of a converging mirror produces an upright, virtual image. Compare this with the situation in Figure 33.54, where a real, inverted image is formed for an object placed outside the mirror's focus. The same occurs with a converging lens: Placing the object inside the focus produces an upright, virtual image (Figure 33.31), while placing the object outside the focus produces a real, inverted image (Figure 33.30).

Diverging mirrors, like diverging lenses, always form virtual images when used alone because the light rays must diverge from the mirror surface. **Figure 33.56** shows a ray diagram for an image formed by a diverging mirror. The image is much smaller than the object, which allows a relatively large scene to be captured on a small mirror surface, and is upright. For these reasons, wide-angle rear-view mirrors on the passenger side of cars and trucks and wide-angle surveillance mirrors are typically convex. (A converging mirror also produces small images of distant objects, but the images are inverted, as Figure 33.54 shows.)

33.19 An object is placed 1.0 m in front of a diverging mirror for which the radius of curvature is 1.0 m. (*a*) Where is the image located relative to the mirror? Is the image real or virtual? (*b*) If the object is 0.30 m tall, what is the height of the image?

33.8 Lensmaker's formula

The focal length of a lens is determined by the refractive index n of the material of which the lens is made and by the radii of curvature R_1 and R_2 of its two surfaces (**Figure 33.57*a***). In this section we work out the relationship among f, R_1, R_2, and n. In this analysis, we can think of a double-convex lens as two plano-convex lenses placed with the two flat surfaces facing each other (Figure 33.57*b*). Remember that both foci of a thin lens are the same distance f from the center of the lens. Because this is true, we can interchange the two surfaces of a thin lens without changing its focal length.

We begin by determining the focal lengths of the two plano-convex lenses in Figure 33.57*b*. **Figure 33.58** on the next page shows a ray diagram for light that passes through the right lens only. To calculate the focal length f_1 of this lens, consider a ray incident from the left that comes in parallel to the axis at a distance h above the axis. Because the ray is normal to the planar surface of the lens, it is not refracted at that surface. After passing through the lens, it strikes the curved surface at an angle θ_i measured from the normal to the curved surface and is refracted as it leaves the lens. The refracted ray emerges at an angle θ_r measured from the normal to the curved surface and crosses the lens axis a distance f_1 from the lens. Therefore the angle that the emerging ray makes with the lens axis is $\theta_r - \theta_i$.

Figure 33.57 Analysis of a double-convex lens.

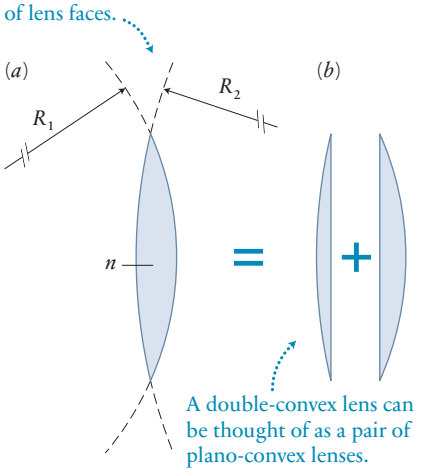
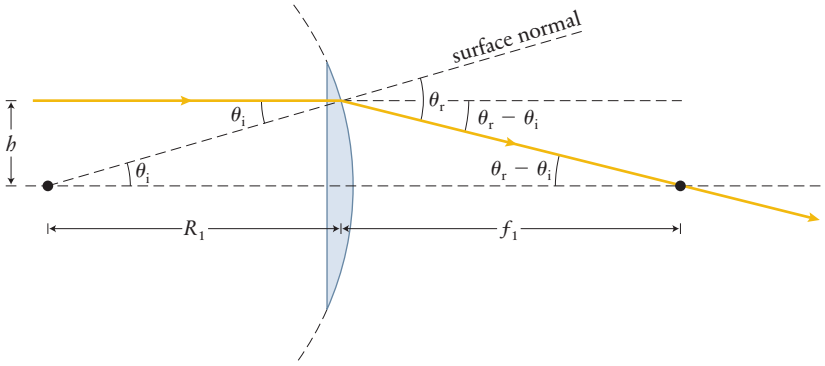


Figure 33.58 Ray diagram for light passing through the right-hand plano-convex lens of Figure 33.57*b*.



Applying Snel's law (Eq. 33.7) to this situation gives

$$n \sin \theta_i = \sin \theta_r. \tag{33.25}$$

(We do not need to show an n on the right because the medium is air and $n_{air} = 1$.) For paraxial rays, we can approximate the sines of the angles in Eq. 33.25 by the angles

$$n\theta_i = \theta_r \quad \text{(small angles)}. \tag{33.26}$$

Using this relationship, we can express the angle between the emerging ray and the lens axis as

$$\theta_r - \theta_i = n\theta_i - \theta_i = (n-1)\theta_i. \tag{33.27}$$

For small angles, the angles are approximately equal to their tangents, which means this relationship can be expressed as

$$\theta_r - \theta_i = (n-1)\theta_i \approx \frac{h}{f_1} \tag{33.28}$$

and

$$\theta_i \approx \frac{h}{R_1}. \tag{33.29}$$

Substituting Eq. 33.29 into Eq. 33.28 and dividing both sides by h, we get

$$\frac{1}{f_1} = \frac{n-1}{R_1}. \tag{33.30}$$

We can follow the same procedure for the left lens in Figure 33.57*b*, using R_2 as our radius of curvature and a ray that originates at the left focus of that lens. The result analogous to Eq. 33.30 is

$$\frac{1}{f_2} = \frac{n-1}{R_2}. \tag{33.31}$$

Now let us determine the focal length of the lens combination by working out the lens equation (Eq. 33.16) for the combination, just as we did for the microscope and telescope in Section 33.6. Consider both the object and light source to be on the left side of the lens combination in Figure 33.57. First, the light from the object strikes the left lens from the left and forms an

image someplace to the right of the right lens. The lens equation that relates the location of this object and image, in terms of the focal length f_2 calculated in Eq. 33.31, is

$$\frac{1}{o_2}+\frac{1}{i_2}=\frac{1}{f_2}=\frac{n-1}{R_2}. \tag{33.32}$$

The image formed by the left lens now serves as the (virtual) object for the right lens. Consequently, $o_1 = -i_2$. (The object for the right lens is virtual, and therefore the object distance o_1 is negative because the object is located on the right side of the lens and the illumination comes from the left side of the lens.) The lens equation for the right lens is thus

$$-\frac{1}{i_2}+\frac{1}{i_1}=\frac{1}{f_1}=\frac{n-1}{R_1}, \tag{33.33}$$

where the rightmost equality comes from Eq. 33.30.

Adding Eqs. 33.32 and 33.33 yields

$$\frac{1}{o_2}+\frac{1}{i_1}=(n-1)\left(\frac{1}{R_1}+\frac{1}{R_2}\right). \tag{33.34}$$

The lens equation for the lens as a whole is simply

$$\frac{1}{o}+\frac{1}{i}=\frac{1}{f}, \tag{33.35}$$

where f, o, and i are the focal length, object distance, and image distance of the lens combination, respectively. The object of the left lens is the actual object, and the image formed by the right lens is the final image, which means that in Eq. 33.35 $o = o_2$ and $i = i_1$. Comparing Eqs. 33.34 and 33.35 gives us the **lensmaker's formula** for the focal length of our lens combination:

$$\frac{1}{f}=(n-1)\left(\frac{1}{R_1}+\frac{1}{R_2}\right). \tag{33.36}$$

Our derivation was for a double-convex lens, but it can be shown that the lensmaker's formula applies to any thin lens, not just a double-convex lens. The radii of curvature are positive for convex surfaces, negative for concave surfaces, and infinity for planar surfaces. For a double-convex lens, f is positive. For a double-concave lens, f is negative (because $n > 1$ for any material used for lenses).



33.20 How should the lensmaker's formula be modified if a lens for which the index of refraction is n_1 is submerged in a medium for which the index of refraction is n_2?

Chapter Summary

Rays (Sections 33.1, 33.2, 33.3, 33.5)

Concepts A **ray** is a line that represents the direction in which light travels.

For **reflected light,** the **angle of incidence** θ_i is the angle between the incident ray and the normal to the reflecting surface, and the **angle of reflection** θ_r is the angle between the reflected ray and the normal to the surface. The **law of reflection** says that when a ray strikes a smooth surface, $\theta_r = \theta_i$ and the two angles are in the same plane.

Refraction is the bending of light when it travels from one medium to another. The **angle of refraction** is the angle between the refracted ray and the normal to the surface at the interface between the materials.

Light passing from a high-density medium into a lower-density medium is bent away from the normal to the surface at the interface between the media. The **critical angle** θ_c is the angle of incidence for which the angle of refraction is 90°, which means the refracted rays emerge along the interface. If the angle of incidence is greater than θ_c, the light is all reflected back into the denser medium (**total internal reflection**).

An **image** is formed when light rays emanating from an object either intersect or appear to intersect at some location in space. The image is *real* when the rays actually do intersect at the location of the image and *virtual* when the rays do not pass through the location of the image.

Dispersion is the spatial separation of waves of different wavelengths in a light ray. When caused by refraction, the separation occurs because the speed at which each wave in the ray moves through a medium depends on the frequency of the wave.

Fermat's principle states that the path taken by a light ray between any two locations is the path for which the time interval needed for the ray to travel between those locations is a minimum.

Quantitative Tools The **index of refraction** n of a medium is

$$n \equiv \frac{c_0}{c}, \tag{33.1}$$

where c_0 is the speed of light in vacuum and c is the speed of light in the medium.

If light has wavelength λ in vacuum, its wavelength λ_1 in a medium having index of refraction n_1 is

$$\lambda_1 = \frac{\lambda}{n_1}. \tag{33.3}$$

Snel's law: When light passes from a medium having index of refraction n_1 into a medium having index of refraction n_2, the angles and indices satisfy the relationship

$$n_1 \sin\theta_1 = n_2 \sin\theta_2, \tag{33.7}$$

where θ_1 is the angle of incidence and θ_2 is the angle of refraction.

For light traveling from a high-density medium having index of refraction n_2 to a lower-density medium having index of refraction $n_1 < n_2$, the **critical angle** is

$$\theta_c = \sin^{-1}\left(\frac{n_1}{n_2}\right). \tag{33.9}$$

Forming images using thin lenses (Sections 33.4, 33.5, 33.6, 33.8)

Concepts A **lens** is an optical element that redirects light rays by refraction such that they form images. The *axis* of a lens is a line that passes through the lens center and is perpendicular to the lens surface. Rays that run close to the lens axis are called *paraxial* rays. They can be either parallel to the lens axis or at a small angle to it.

Quantitative Tools The **lens equation** is

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i}, \tag{33.16}$$

where f is the **focal length** of the thin lens, o is the *object distance*, and i is the *image distance*. The sign possibilities for f, o, and i are given in Table 33.2.

Rays that are parallel to and close to the axis of a thin *converging* (*convex*) lens converge to a point called the **focal point** (or **focus**) of the lens. For a thin *diverging* (*concave*) lens, rays entering the lens from the left parallel to and close to the axis diverge as though they came from a single point to the left of the lens, called the **focal point.**

The **focal length** of a lens is the distance from the focal point to the center of the lens.

To determine the location and orientation of an image formed by a lens, make a ray diagram using the **principal rays.** See the accompanying Procedure boxes.

The **magnification** of a lens is

$$M \equiv \frac{h'}{h} = -\frac{i}{o}, \tag{33.17}$$

where h' is the image height and h is the object height. The image is upright if M is positive and inverted if M is negative.

Lensmaker's formula: The focal length of a lens made of material of index of refraction n, having radii of curvature R_1 and R_2, and surrounded by air is given by

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right). \tag{33.36}$$

The radii are positive for convex surfaces, negative for concave surfaces, and infinity for planar surfaces.

Table 33.2 Sign conventions for f, i, and o (positive = real; negative = virtual)

Sign	Lens	Mirror
$f > 0$	converging lens	converging mirror
$f < 0$	diverging lens	diverging mirror
$o > 0$	object in front[b] of lens	object in front of mirror
$o < 0$[a]	object behind lens	object behind mirror
$i > 0$	image behind lens	image in front of mirror
$i < 0$	image in front of lens	image behind mirror
$h_i > 0$	image upright	image upright
$h_i < 0$	image inverted	image inverted
$\|M\| > 1$	image larger than object	image larger than object
$\|M\| < 1$	image smaller than object	image smaller than object

[a]Encountered only with lens or mirror combinations.
[b]For both lenses and mirrors, *in front* means *on the side where the rays originate; behind* refers to the opposite side.

Optical instruments (Section 33.6)

Concepts The lens of the eye forms an image on the retina. The *near point* is the closest object distance at which the eye can focus comfortably, about 0.25 m from the eye for an adult.

A compound microscope uses two converging lenses to form its image. The image formed by the first lens (the *objective* lens) serves as the object for the second lens (the *eyepiece* lens).

A refracting telescope uses two converging lenses to form its image. They are positioned so that the objective lens forms a real image of a distant object and the eyepiece lens forms, at infinity, a virtual image of this real image.

Quantitative Tools The **angular magnification** M_θ of a lens is

$$M_\theta = \left|\frac{\theta_i}{\theta_o}\right|, \tag{33.18}$$

where θ_i is the angle subtended by the image and θ_o is the angle subtended by the object.

For small angles, the angular magnification of a converging lens placed between an object and the eye is

$$M_\theta \approx \frac{0.25\ \text{m}}{f}. \tag{33.21}$$

The *strength* d of a lens measured in diopters is

$$d \equiv \frac{1\ \text{m}}{f}. \tag{33.22}$$

The magnification M of a compound microscope is

$$M = M_1 M_{\theta 2} = \frac{-0.25\ \text{m}}{f_2\left(\frac{o_1}{f_1} - 1\right)},$$

where M_1 is the magnification of the objective lens, $M_{\theta 2}$ is the angular magnification of the eyepiece lens, o_1 is the distance from the object to the objective lens, f_1 is the focal length of the objective lens, and f_2 is the focal length of the eyepiece lens.

The angular magnification M_θ of a refracting telescope is

$$M_\theta = \left|\frac{\theta_i}{\theta_o}\right| \approx \left|\frac{f_1}{f_2}\right|,$$

where f_1 is the focal length of the objective lens and f_2 is the focal length of the eyepiece lens.

Mirrors (Sections 33.2, 33.7)

Concepts A flat mirror forms a virtual image of an object. The image is just as far behind the mirror as the object is in front of the mirror.

A *concave* spherical mirror focuses incident rays that are parallel to and close to the axis of the mirror to a point in front of the mirror (the **focal point**), corresponding to a real focus. A *convex* spherical mirror disperses the parallel incident rays such that they appear to come from a point behind the mirror (the **focal point**), corresponding to a virtual focus.

To determine the location and orientation of an image formed by a mirror, make a ray diagram showing the three **principal rays.** See the accompanying Procedure boxes.

Quantitative Tools The focal length of a spherical mirror is one-half its *radius of curvature*:

$$f = \frac{R}{2}. \tag{33.23}$$

For a spherical mirror, the relationship of focal length f, object distance o, and image distance i is given by the lens equation:

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i}. \tag{33.24}$$

The sign possibilities for f, o, and i are given in Table 33.2.

Questions and Problems

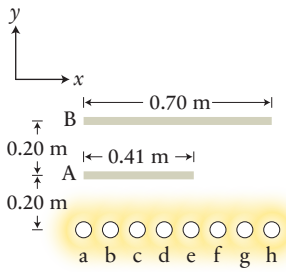
For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy,* •• = *intermediate,* ••• = *hard;* CR = *context-rich problem.*

33.1 Rays

1. Figure P33.1 shows eight point sources of light, a–h, and two detectors, A and B. The point sources are evenly spaced, with 0.10 m between adjacent sources, and detector A shields part of detector B from the point sources. If no light is reflected from any nearby objects, how many of the point sources illuminate any part of (*a*) detector A and (*b*) detector B? •

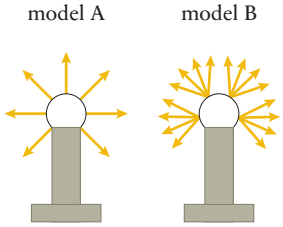
Figure P33.1



2. A point source is fixed 1.0 m away from a large screen. Call the line normal to the screen surface and passing through the center of the point source the z axis. When a sheet of cardboard in which a square hole 0.040 m on a side has been cut is placed between the point source and the screen, 0.50 m from the point source with the hole centered on the z axis, a bright square shows up on the screen. If, instead, a second sheet of cardboard with a similar square hole is placed between the point source and screen, 0.25 m from the point source with the hole centered on the z axis, the bright square it casts on the screen is identical to the bright square from the first sheet. What are the dimensions of the hole in this sheet? •

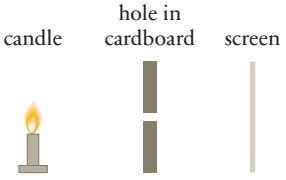
3. A thin, flat object in which a square hole 30.0 mm on a side has been cut is illuminated with a point source. Light that passes through the hole strikes a screen 300 mm behind the object. Call the direction normal to the screen the z direction. The point source is shifted away from the center of the hole by 150 mm in the positive x direction. The point source is centered on the hole in the y direction and is 300 mm in front of the object in the z direction. What are (*a*) the dimensions of the bright region on the screen and (*b*) the x and y coordinates of the perimeter of the bright region? ●●
4. Sunlight that reaches Earth is blocked by Earth from traveling any farther, so that the region on the side of Earth facing away from the Sun is in shadow. The portion of this region in complete shadow is the *umbra*, and the portion in partial shadow is the *penumbra*. The Sun's radius is 7.0×10^5 km, Earth's radius is 6.4×10^3 km, and the Sun–Earth distance is 1.5×10^8 km. (*a*) What is the shape of the umbra? Of the penumbra? (*b*) Consider an observer in space on the side of Earth facing away from the Sun—that is, in the shadow region. Does this observer see the Sun when he is in the umbra? When he is in the penumbra? ●●
5. Two models of light emitted from a light bulb are illustrated in Figure P33.5. (*a*) Describe the difference in the behavior of light in each model. (*b*) Describe an experiment that can determine which model of light is more accurate. ●●

Figure P33.5



6. An aquarium light bulb has a long, straight glowing filament surrounded by a tubular, clear glass bulb. This lighted bulb, with its filament oriented vertically, is placed 1.0 m from a screen in an otherwise dark room. A thin sheet of cardboard is placed halfway between the filament and the screen so that no light can reach the screen. If a small hole, shaped like an equilateral triangle, is cut through the center of the cardboard sheet, describe the pattern of light that appears on the screen. ●●●
7. A burning candle is placed in front of a screen, and a sheet of cardboard with a small round hole in it is inserted between the candle and the screen as shown in Figure P33.7. Describe the pattern of light that appears on the screen. ●●●

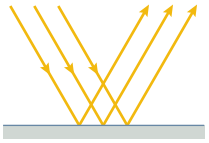
Figure P33.7



33.2 Absorption, transmission, and reflection

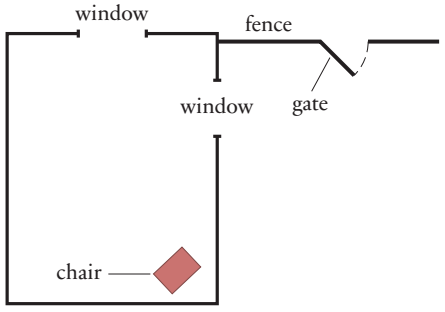
8. Figure P33.8 shows light rays reflected from a smooth surface. Draw three wavefronts along the rays that have not yet struck the surface and three wavefronts along the reflected rays. ●

Figure P33.8



9. For the room shown in Figure P33.9, draw a diagram that shows where a mirror should be mounted so that a person sitting in the chair has a full view of the gate in the fence outside. ●

Figure P33.9



10. A classmate exclaims about the professor's bright blue shirt, "Wow, that shirt has a lot of blue in it!" Evaluate your classmate's statement. ●
11. A flat mirror oriented perpendicular to the y axis of a Cartesian coordinate system extends along the x axis from $x = 0$ to $x = 4.0$ m. An object is located in the xy plane at the point (1.0 m, 2.0 m). With your eye in the xy plane in the range $y < 0$ and x equal to any value, you cannot see the image of this object in the mirror. For what range of x and y in the xy plane can you see the image in the mirror? ●●
12. The bottom edge of a wall-mounted mirror aligns with your waist, and the top edge aligns with a point somewhere above your head. Draw a diagram that illustrates, when you stand in front of this mirror, the portion of your whole body that you can see in the image. ●●
13. The minimum height of a wall mirror in which you can view your entire body is half your height, with the distance from the floor to the bottom of the mirror equal to half the distance between your eyes and feet. This relationship between mirror height/position and object height does not hold for objects that are far behind you, however. How large (height, width, or both) a reflecting surface is needed in order for you to see the

entire image of a large distant object—a tree, say? (Hint: What quantity is different when you view your own image as opposed to when you view the image of a distant object?) ●●

14. Standing 1.0 m away from a large mirror, you see an image of a flashlight in the mirror. If the image appears to be 2.0 m behind the mirror, 1.0 m vertically above your eyes, and 3.0 m to your right (your actual right, not your reflection's right), what is the flashlight's location relative to you? ●●

15. A classmate believes that the image you see in a mirror is on the mirror surface rather than behind the mirror. What is something you can do to convince your classmate he is incorrect? ●●

16. A woman holds a mirror stationary and perfectly vertical as a boy runs toward it at a speed of 1.0 m/s. (*a*) In her reference frame, what does the woman observe for the speed at which the boy and his reflection move toward the mirror? (*b*) In the reference frame of the boy, what does he observe for the speed at which the woman and his reflection move toward him? ●●

17. Consider this statement: Think of the reflection of the Moon on a cloudless night: Look in the ocean, you see a large moon. Look in a pond; the same moon is there. In a puddle, in a teacup—everywhere, diminished somewhat in size, but all the features are there. (*a*) What physical attribute (referred to only indirectly in the statement) is it that differs in size for the ocean, pond, puddle, and cup? (*b*) What other quantity or quantities determine this attribute? Draw a diagram showing how the image of the Moon forms and relating all the quantities you've identified to one another. ●●●

18. You want to see how your new belt buckle matches your new hat. Your eyes are 110 mm below the top of the hat and 800 mm above the buckle. You walk up to your mirror and notice that the mirror is exactly large enough and exactly in the right position for you to see both the top of the hat and the buckle. (*a*) What is the position of the top edge of the mirror relative to your eyes? (*b*) How tall is the mirror? ●●●

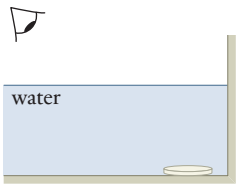
33.3 Refraction and dispersion

19. The light bulb in a bathroom ceiling is directly over a toy boat floating in a bathtub filled with water to a depth x. How does the shadow the boat casts on the bottom of the tub compare with the shadow it casts when the water is drained out and the boat is held the same distance x above the tub bottom? Treat the bulb as a point source. ●

20. While scuba diving, you have injured yourself and must signal the boat on the surface for help. You have five laser pointers, each emitting light of a different color: red, orange, yellow, green, and blue. When you shine the yellow light up toward the boat, the beam makes such a small angle with the surface that all the light is reflected back into the water. (*a*) If you are not able to move closer to the boat, which color or colors of light might you try? (*b*) Are these other lights likely to be successful in penetrating the surface at that point? ●

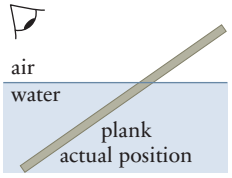
21. The observer in Figure P33.21 is looking at a coin near the edge of a pool filled with water. To this observer, does the water appear to be shallower or deeper than it really is? ●

Figure P33.21



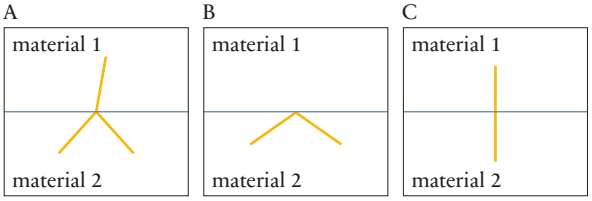
22. In Figure P33.22, a straight plank of lumber enters a pool of water at an angle. What does the plank look like to the observer? ●●

Figure P33.22



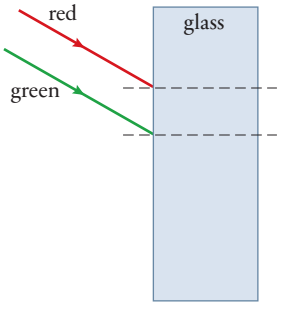
23. Figure P33.23 depicts incident and all reflected or refracted light rays in three cases, A, B, and C. For each case, state (*a*) in which material, 1 or 2, the light travels faster and (*b*) in which material, 1 or 2, the ray originates. ●●

Figure P33.23



24. Parallel red and green laser rays are incident on a glass slab as shown in Figure P33.24. Sketch the rays as they pass through the slab and after they have entered the air to the right of the slab. Are the rays parallel after they exit the slab? Is the distance between the rays after they exit the slab equal to the distance between them before they entered? ●●

Figure P33.24

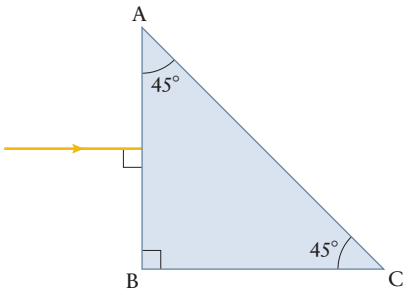


25. A thick layer of transparent oil floats on top of water. Sketch what happens to a ray of light coming in at a 30° angle away from the normal to the surface. ●●

26. Suppose a small flashlight bulb is on the bottom of the bathtub of Problem 19, directly under the toy boat. When this bulb is lit and the ceiling light is turned off, how does the size of the shadow cast by the boat on the ceiling change when the water in the tub is replaced by a liquid in which the speed of light is (*a*) faster than the speed of light in water and (*b*) slower than the speed of light in water? ●●

27. Suppose the prism in Figure P33.27 is immersed in a liquid in which the speed of light is lower than the speed of light in glass. Describe what happens to the light shown entering at normal incidence. ••

Figure P33.27



28. Driving down a long straight highway, you see a reflection of the sky in the road ahead. It disappears before you get to it but reappears down the road continually. Because there is nothing different about the road from one location to another, explain what is happening in terms of refraction. ••
29. White light passing through a prism is dispersed into its constituent colors. Describe how this rainbow of colors can be changed back to white light. •••
30. A glass cylinder is placed inside a large beaker, and corn oil is poured into the combination until the beaker is full. If the glass cylinder is now invisible inside the oil, what must you conclude about how light interacts with the oil and the glass? •••

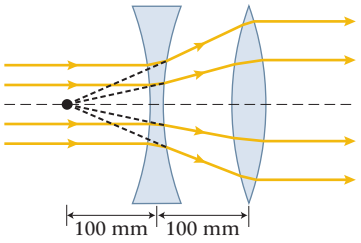
33.4 Forming images

31. (*a*) Draw a simplified ray diagram showing the three principal rays for an object located outside the focal length of a converging lens. (*b*) Is the image real or virtual? (*c*) Is it upright or inverted? (*d*) What happens to the image as the object is moved farther from the lens? •
32. (*a*) Draw a simplified ray diagram showing the three principal rays for an object located outside the focal length of a diverging lens. (*b*) Is the image real or virtual? (*c*) Is it upright or inverted? (*d*) What happens to the image as the object is moved farther from the lens? •
33. The lens in a magnifying glass is a converging lens, but parallel light rays that pass through a converging lens must converge to a point. This being true, how can a magnifying glass form an image that is larger than the object? Why isn't the image reduced to a point? •
34. A lens has a radius of curvature of magnitude R on its left surface and a radius of curvature of magnitude $2R$ on its right surface. Parallel rays entering the lens from the left are focused to a point 100 mm to the right of the right surface. At what distance from the lens is the point at which parallel rays entering from the right are focused? •
35. Looking at an object through a thin lens, you see a virtual image. Which side of the lens is the image on, or does it appear to be inside the lens? What property of this image tells you whether the lens is converging or diverging? ••
36. (*a*) Draw a simplified ray diagram showing the three principal rays for an object located inside the focal length of a diverging lens. (*b*) Is the image real or virtual? (*c*) Is it upright or inverted? (*d*) What happens to the image as the object is moved toward the focal point? ••
37. (*a*) Draw a simplified ray diagram showing the three principal rays for an object located inside the focal length of a converging lens, closer to the lens than to the focal point. (*b*) Is the image real or virtual? (*c*) Is it upright or inverted? (*d*) What happens to the image as the object is moved closer to the focal point? ••
38. An object 20 mm tall is placed 70 mm to the left of a converging lens that has a focal length of 100 mm. (*a*) Draw a scale diagram showing the three principal rays, and use them to locate the image. Is the image (*b*) real or virtual, (*c*) inverted or upright, (*d*) larger or smaller than the object? ••
39. An object is placed 80 mm to the left of a diverging lens for which the focus is 100 mm from the lens. (*a*) Draw a scale diagram showing the three principal rays, and use them to locate the image. Is the image (*b*) real or virtual, (*c*) inverted or upright, (*d*) larger or smaller than the object? ••
40. Complete the table below for a converging lens. ••

		Image		
Object Location	**Image Location**	**Real or virtual?**	**Upright or inverted?**	**Larger or smaller than the object?**
Between lens and focus	Between positive infinity and lens (same side)	Virtual	Upright	Larger
At focus				
Between focus and twice the focal length				
At twice the focal length				
Beyond twice the focal length				
At infinity				

41. Figure P33.41 shows light passing first through a diverging lens and then through a converging lens placed 100 mm to the right of the diverging lens. The rays emerge from the converging lens parallel to one another. Calculate the distance from each lens to its focus. ••

Figure P33.41

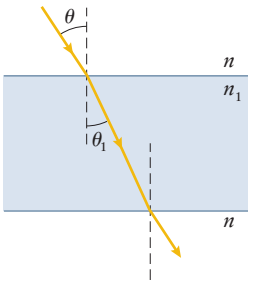


42. Two identical converging lenses of focal length 100 mm are placed 400 mm apart, and an object 40 mm tall is placed 230 mm to the left of the left lens. (*a*) Draw a scale diagram showing the three principal rays for the left lens, and use them to locate the image formed by this lens. (*b*) Draw the principal rays for the right lens, and use them to locate the image formed by this lens. Is the image formed by the right lens (*c*) real or virtual, (*d*) inverted or upright, (*e*) larger or smaller than the object? •••

33.5 Snel's law

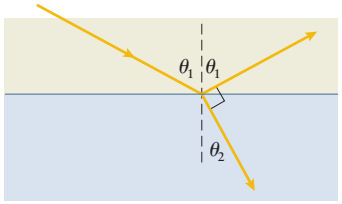
43. Light is traveling at a speed of 1.82×10^8 m/s through a certain material. Determine what material it is likely traveling through. •
44. What is the index of refraction of a material in which the speed of light is 20% less than the speed of light in vacuum? •
45. A red laser light has a wavelength of 700 nm in air. What are the frequency and wavelength of this light in flint glass (index of refraction $n = 1.65$)? •
46. A light ray traveling through medium 1 with index of refraction $n_1 = 1.70$ reaches the interface between medium 1 and medium 2 with index of refraction $n_2 = 1.50$. (*a*) At what minimum angle with respect to the normal must the ray be incident on the interface in order to be totally internally reflected? (*b*) How would the situation change if the ray were traveling in medium 2 when it reached the interface? •
47. A light ray travels up from the bottom of a tank containing oil (index of refraction $n = 1.48$). If the oil surface is smooth, at what minimum angle from the normal to that surface would the light ray have to strike the surface in order to reflect back into the tank and not escape? •
48. You shine a laser pointer at an angle of 30° from the normal onto the surface of a thick block of glass for which the index of refraction is $n = 1.5$. (*a*) What is the angle of refraction of the laser beam? (*b*) Draw a diagram showing the normal, the surface of the block, and the incident and refracted wavefronts and rays. (*c*) How would the angles in parts *a* and *b* change if the incident angle were 45°? ••
49. In vacuum, the wavelength of the red light in your laser pointer is 680 nm. Standing at the shore of a calm pond, you shine the laser beam onto the water surface at an angle of 45° from the normal to the surface. If the index of refraction for the water is $n = 1.333$, how do each of the following properties of the light change as the light enters the water: (*a*) wavelength, (*b*) frequency, (*c*) speed, and (*d*) direction of propagation? (*e*) If you shine the beam onto the surface along the normal, do any of your answers change and, if so, how? ••
50. A light ray is incident on a layer of oil floating on water. The index of refraction of oil is greater than that of water, which is greater than that of air. Call the angle of incidence in air θ_a, the angle of refraction in oil θ_o, and the angle of refraction in water θ_w. (*a*) Draw a diagram showing all three layers and all angles. (*b*) Show that $n_a \sin\theta_a = n_w \sin\theta_w$. (*c*) The result in part *b* implies that in calculating the final direction (θ_w) of a light ray passing from medium 1 into medium 2 and then into medium 3, all having different n values, we can ignore the presence of medium 2. Can you also ignore medium 2 if what you are calculating is the *location* where the ray strikes the bottom of the container of water? ••
51. The index of refraction in flint glass is 1.66 for blue light and 1.61 for red light, with the values for other colors between these limits. A ray of white light traveling from left to right through a slab of flint glass undergoes dispersion as it exits the slab and enters the air. If the exiting ray strikes the right surface of the slab at an angle of 30.0° from the normal to the surface, what is the angle of refraction (*a*) for the blue light and (*b*) for the red light? (*c*) A flat opaque screen is placed in the path of the refracted rays such that the red ray strikes perpendicular to the screen. If the distance along the refracted red ray from slab to screen is 0.50 m, what is the distance between the point of red light and the point of blue light in the rainbow produced on the screen? (Use the small-angle approximation.) ••
52. What is the critical angle for total internal reflection for an optical fiber cable surrounded by air if the index of refraction is (*a*) 1.4 and (*b*) 1.8? (*c*) Which fiber could have a sharper bend (smaller radius of curvature) and still be capable of total internal reflection? ••
53. Show that the ray exiting the block in Figure P33.53 is parallel to the ray entering the block. ••

Figure P33.53



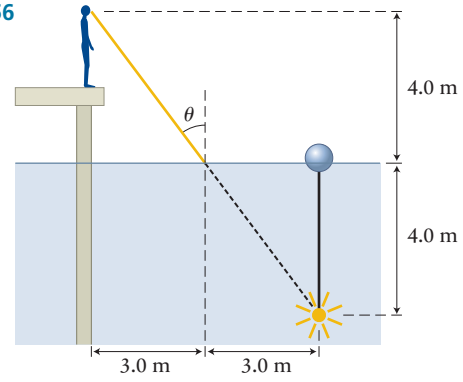
54. Green laser light for which the wavelength in air is 550 nm strikes a flint glass ($n_{glass} = 1.65$) at an angle of 50.0° from the normal to the air-flint glass interface. Calculate (*a*) the wavelength and frequency of the light inside the flit glass and (*b*) the angle of refraction. ••
55. A light ray travels through medium 1, index of refraction $n_1 = 1.5$, into medium 2, index of refraction $n_2 = 2.4$ (Figure P33.55). At what angle with respect to the normal to the interface between the media does the light need to strike the interface so that the reflected and the refracted rays are at right angles to each other? ••

Figure P33.55



56. A researcher stands at the end of a lakeside pier, 4.0 m above the water surface and 6.0 m away from a buoy floating on the lake (Figure P33.56). A rope hangs from the bottom of the buoy down into the water, with a light attached to its end. When the researcher looks at the light from a position at which his line of sight intersects the water surface exactly halfway between the pier and the buoy, the light appears to be 4.0 m below the water surface. How far under the surface is the light? ●●

Figure P33.56



57. Figure P33.57 depicts a light ray entering an optical fiber from air and being transmitted through the fiber by total internal reflection. (*a*) If the index of refraction of the fiber is n, what is the maximum value of θ, the angle the entering ray makes with the normal to the flat end of the fiber, for which total internal reflection occurs? (*b*) When n is greater than a certain value n_0, the light ray can enter the fiber at any angle θ between 0 and 90° and total internal reflection still occurs inside the fiber. What is n_0? ●●●

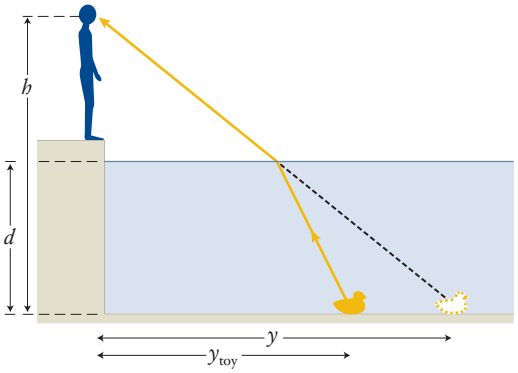
Figure P33.57



58. A ray traveling through a large vat filled with medium 1, index of refraction $n_1 = 1.3$, is incident on a flat layer (thickness $d = 20$ mm) of medium 2, index of refraction $n_2 = 2.4$. If the light enters medium 2 at an angle $\theta_1 = 30°$ with respect to the normal to the interface between the media, how far is it shifted from its original path once it exits medium 2? ●●●

59. A child's toy sits on the bottom of a swimming pool in which the water ($n = 1.33$) depth is $d = 2.0$ m (Figure P33.59). To a child standing at the pool edge and a distance $h = 3.6$ m above the pool bottom, the toy appears to be a horizontal distance $y = 4.8$ m away from the sidewall. At what horizontal distance from the sidewall is the toy located? ●●●

Figure P33.59



33.6 Thin lenses and optical instruments

60. What is the lens strength of a diverging lens that has a focal length of magnitude 200 mm? ●

61. Mathematically, why can't a single diverging lens produce a real image? ●

62. What is the focal length of eyeglasses that have a lens strength of +2.5 diopters? ●

63. Using a converging lens that has focal length $f = 200$ mm, how far behind the lens do you expect to see a sharp image of an object located 1.00 m in front of the lens? ●

64. A cheap cell phone camera uses a single lens to form an image on a sensor that is 10 mm high and 5.8 mm behind the lens. How far do you need to be away from the Eiffel Tower (height 324 m) to get its whole length in a photograph? Ignore the tilting that occurs as you take the photo from the ground. ●

65. The objective lens in a telescope that produces a 50-fold angular magnification has a focal length of 3.0 m. What is the focal length of the eyepiece lens? ●

66. An object is placed 800 mm away from a thin converging lens that has a focal length of 500 mm. What are (*a*) the image distance and (*b*) the magnification? (*c*) Draw a simplified ray diagram for this case. ●●

67. An object is placed 150 mm away from a converging thin lens that has a focal length of 400 mm. What are (*a*) the image distance and (*b*) the magnification? (*c*) Draw a simplified ray diagram for this case. ●●

68. An object is placed 500 mm away from a diverging thin lens for which the focal length is −200 mm. What are (*a*) the image distance and (*b*) the magnification of the image? ●●

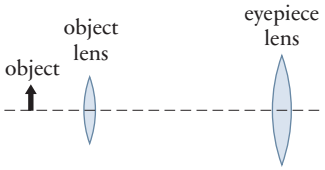
69. An object is placed 200 mm from a diverging thin lens that has a focal length of −500 mm. What are (*a*) the image distance and (*b*) the magnification? (*c*) Draw a simplified ray diagram for this case. ●●

70. A bug is 30 mm away from a lens. As you look through the lens, the bug appears to be just at your near point (0.250 m in front of your eyes). If the actual distance between your eyes and the bug is 130 mm, what is the lens focal length? ●●

71. A woman is not able to focus on objects that are closer to her eyes than 400 mm unless she wears eyeglasses. Her glasses enable her to clearly see objects at the near point of an average human (and no closer). What is the strength of the lenses in her glasses? ●●

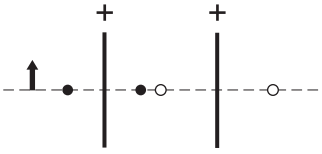
72. Two converging lenses are placed facing each other and 600 mm apart, with lens 1 to the left of lens 2. An object is placed 400 mm to the left of lens 1, which has a focal length of 150 mm. (*a*) If the focal length of lens 2 is 200 mm, where is the image located? (*b*) What is the overall magnification of the image? (*c*) Is the image upright or inverted? ••
73. An astronomical telescope of angular magnification M_θ consists of an objective lens and an eyepiece lens separated by a distance d, such that the focal points of the lenses coincide inside the telescope. In terms of M_θ and d, what are (*a*) the focal length of the objective lens and (*b*) the focal length of the eyepiece lens? ••
74. A schematic diagram of a refractive microscope is shown in Figure P33.74. The focal length of the objective lens is 25 mm, that of the eyepiece lens is 63 mm, and the two lenses are 200 mm apart. If the object is 30 mm away from the objective lens, what are (*a*) the distance between the eyepiece lens and the image produced by that lens and (*b*) the overall magnification of this image? ••

Figure P33.74



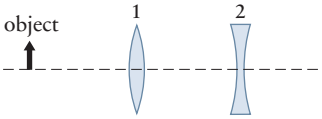
75. Two converging lenses, 1 and 2, are positioned facing each other, with lens 1 on the left. The focal length of lens 1 is 100 mm, that of lens 2 is 180 mm, and the two lenses are 150 mm apart. If an object is 50 mm to the left of lens 1, (*a*) at what distance from lens 2 should a screen be placed if the object's image is to be focused on the screen? (*b*) What is the overall magnification of the image? (*c*) Is the image upright or inverted? ••
76. Two lenses with focal lengths $f_1 = 200$ mm and $f_2 = 300$ mm are placed facing each other a distance $d = 1.5$ m apart, with lens 1 to the left of lens 2. What is the location of the image formed of an object placed 250 mm to the left of lens 1? ••
77. The focal points of the two converging lenses shown in Figure P33.77 are denoted by solid dots for the left lens and open dots for the right lens. Describe the final image. ••

Figure P33.77



78. In Figure P33.78, a converging lens 1, of focal length 100 mm, and a diverging lens 2, of focal length –80.0 mm, are 160 mm apart. If the object is 180 mm away from lens 1, what are (*a*) the distance between lens 2 and the image produced by that lens and (*b*) the overall magnification of this image? •••

Figure P33.78



79. Problem 42 describes a setup of two identical converging lenses spaced 400 mm apart. Each lens has a focal length of 100 mm, and a 40-mm-tall object stands 230 mm to the left of the left lens. (*a*) Use an analytical approach to determine the position of the final image. Based on your calculation, is the final image (*b*) inverted or upright and (*c*) real or virtual? (*d*) Calculate the overall magnification. •••
80. You need to design a magnification system that provides a magnification of 3.00 and creates a real, upright image located 300 mm from the surface of the final lens in the system. The only lenses you have to work with are two converging lenses, each having a focal length of 100 mm. How should you place them with respect to the object, as well as with respect to each other, so as to achieve your goals? •••
81. You are trying to look at a geosynchronous communications satellite that is 5.500 m long. A typical human eye can resolve objects that are 0.1000 mm wide when the objects are at the person's near point, and the same angular size applies to objects farther away than the near point. Suppose the telescope you are using is 1.000 m long from one lens to the other. Describe what must be true of the lenses in order for you to be able to look at the satellite. Assume the satellite is exactly resolvable when viewed through your telescope. Is your result reasonable? What modifications could you make to the telescope in order to make it easier to see the satellite? •••

33.7 Spherical mirrors

82. Sunlight reflected off a mirror you are holding is focused at a point 100 mm in front of the mirror. What is the radius of curvature of the mirror? •
83. A spherical mirror with radius of curvature $R = 200$ mm is used to form an image of an object placed a distance $d = 150$ mm in front of the mirror. Where does the image appear? •
84. A bathroom shaving mirror has radius of curvature $R = 400$ mm. Describe the image of your face formed when you stand (*a*) a distance $d_{close} = 100$ mm from the mirror and (*b*) a distance $d_{far} = 1.20$ m from the mirror. •
85. Passenger-side car mirrors often carry the warning "Objects in mirror are closer than they appear." If the mirror on your car has radius of curvature $R = -800$ mm and the car behind you and one lane to the right is a distance $o = 20.0$ m away, how far away does it appear? •
86. An object is placed 60 mm in front of a converging spherical mirror for which the radius of curvature is 200 mm. What are (*a*) the image distance and (*b*) the magnification of the image? (*c*) Is the image real or virtual? (*d*) Is it upright or inverted? (*e*) Draw a ray diagram illustrating this situation and including the three principal rays. ••
87. A converging mirror that has a radius of curvature of 70.0 mm forms an image of an object that is 20.0 mm tall and 150 mm in front of the mirror. (*a*) What is the image height? (*b*) Is the image real or virtual? (*c*) Is it upright or inverted? (*d*) Draw a ray diagram illustrating this situation and including the principal rays. ••
88. You stand 0.50 m in front of a diverging mirror for which the radius of curvature is −3.5 m. (*a*) Is your image real or virtual? (*b*) Is it upright or inverted? (*c*) How far away from the mirror surface is it located? (*d*) What is its magnification? ••

89. When you look at yourself in a converging mirror, you appear to be one-fourth your actual size. If you are standing 1.0 m in front of the mirror, what is the mirror's radius of curvature? ••

90. You wish to use a converging mirror to form an image that is N times larger than the object. In terms of N and R, the radius of curvature of the mirror, at what distance o from the mirror should you place the object? ••

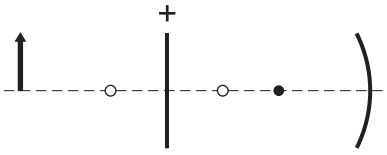
91. A spherical mirror produces a real image of your face that is twice as large as your face. If your face is 750 mm from the mirror, what is the radius of curvature of the mirror? ••

92. A converging mirror has a focal length of 200 mm. Calculate the image distance and magnification for an object located (*a*) halfway between the focal point and the mirror, (*b*) at the focal point, (*c*) halfway between the focal point and the center of curvature, (*d*) at the center of curvature, (*e*) a distance f beyond the center of curvature, and (*f*) at infinity. ••

93. Repeat Problem 92 for a diverging mirror. Comment on any similarities or differences with the two types of mirrors. ••

94. A security mirror in a shop is designed to give a wide-angle view, but as a result the images it forms are much smaller than their objects. If you want an image that appears 10.0 m behind the mirror to have a magnification of 10.0%, what radius of curvature must the mirror have? ••

Figure P33.95



95. The optical system of Figure P33.95 consists of a converging lens and a converging mirror. The focal points of the lens are marked by open dots and the center of curvature of the mirror by a solid dot. Draw a simplified ray diagram, and describe the final image. •••

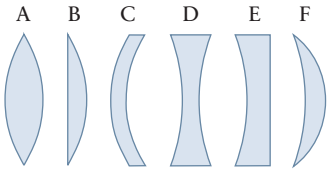
96. The kind of image a spherical mirror produces for a real object depends on where the object is located relative to the focal point of the mirror. There are three types of spherical mirrors—converging, diverging, and flat—if you consider a flat mirror as being one portion of a curved surface for which the radius of curvature is infinitely large. For a real object placed in front of each type of mirror, characterize the image as real or virtual; upright or inverted; in front of or behind the mirror; and enlarged, shrunken, or the same size as the object. •••

97. You position an object in front of a spherical mirror for which the magnitude of the radius of curvature is R_1. The image formed is real, inverted, and twice as large as the object. You then replace this spherical mirror with one that has a radius of curvature of magnitude R_2 (keeping the object distance unchanged). The image formed by this mirror is virtual, upright, and twice as large as the object. Calculate a value for the ratio R_2/R_1. •••

33.8 Lensmaker's formula

98. All the lenses in Figure P33.98 are surrounded by air. Which of the lenses are converging, and which are diverging? •

Figure P33.98



99. You have a thin lens that is convex on both surfaces. The material of which the lens is made has index of refraction $n = 1.40$, and the radii of curvature are $|R_1| = 300$ mm and $|R_2| = 500$ mm. What is the lens focal length in air? •

100. To correct the vision of a near-sighted patient, an optician needs to make a pair of eyeglasses using –3.0-diopter lenses. The lenses are plano-concave, and the front surface (that is, the surface farthest from the eye) is the flat surface. If the glass used for the lenses has an index of refraction of 1.5, what radius of curvature must the remaining surface have? •

101. A plano-convex lens has a focal length of 150 mm, and the material out of which the lens is made has an index of refraction of 1.5. What is the radius of curvature of the convex surface? •

102. The radius of curvature of one surface of a glass lens ($n = 1.45$) is 1.50 m. If sunlight is focused at a point 0.300 m away from the lens, what is the radius of curvature of the other surface? ••

103. The two surfaces of a double convex, thin lens made of plastic each have an absolute value of radius of curvature of 1.8 mm. If the focal length of the lens is 4.5 mm, what is the index of refraction of the plastic? ••

104. A lens has an index of refraction of 1.50. Its left side curves outward with an absolute value of radius of curvature of 0.15 m, and its right side curves inward with an absolute value of radius of curvature of 0.25 m. What is its focal length if you look through it (*a*) from the left side and (*b*) from the right side? ••

105. A thin lens is made of material that has an index of refraction of 1.50. When an object 10 mm tall is placed 500 mm away from the lens, an upright image 21.5 mm tall is formed. One surface of the lens is concave, with an absolute value of radius of curvature of 350 mm. What is the radius of curvature of the other surface? ••

106. A lens made of glass for which the index of refraction is 1.55 has a focal length of 0.500 m in air. What is its focal length when this lens is submerged in water? ••

107. The convex surface of a plano-convex lens has an absolute value of radius of curvature of 40 mm. The index of refraction for the glass of which the lens is made is 1.5. (*a*) What is the focal length of the lens? (*b*) You take some measurements and determine that the 40-mm radius of curvature is correct but the focal length is 100 mm instead of the value calculated in part *a*. What must the index of refraction for the glass be? ••

108. A thin lens of focal length f_A is made of a material for which the index of refraction is $n_A = 1.1$. If you make a lens B that is identical to lens A in every way except that the index of refraction of the material used to make B is $n_B = 2n_A$, how does the focal length f_B for this lens compare with f_A? ••

109. The thin lens in air of Figure P33.109 is drawn showing $|R_1| < |R_2|$. (*a*) Is this lens converging or diverging? (*b*) If the lens were drawn showing $|R_1| = |R_2|$, would it be converging or diverging? (*c*) If it were drawn showing $|R_1| > |R_2|$, would it be converging or diverging? ●●

Figure P33.109



110. The two surfaces of a lens made of an unknown material have the same magnitude of radius of curvature $|R|$. When an object is placed a distance R away from the lens, the image formed is real and twice as large as the object. What is the index of refraction of the material of which the lens is made? ●●●

Additional Problems

111. When you shine a laser-beam pointer at a wall, why do you see a dot on the wall but not the beam running from pointer to wall (in the absence of dust or mist)? ●

112. A man spearfishing from a dock sees a fish in the water. In order to hit the fish, should he aim the speargun above, below, or directly at the image he sees? ●

113. Traveling through an optical cable for which the index of refraction is 1.6, how long does it take a light signal to travel from New York to Los Angeles, a distance of 4.0×10^3 km? ●

114. In about the years 214–212 BC, the Greek philosopher and scientist Archimedes is said to have repelled a Roman attack with a "death-ray" converging mirror, using sunlight to set the approaching fleet on fire. Assume an early morning attack from the east, with the Sun directly behind the ships. If the approaching fleet is 100 m away from the mirror, what would the mirror's radius of curvature have to be in order to achieve the maximum effect? ●

115. What is an application of a converging lens for these situations: (*a*) object is at twice the focal length, (*b*) object is at the focal length, and (*c*) object is at infinity? ●●

116. An object that is 50 mm tall is placed 75 mm in front of a converging thin lens. If the image is inverted and 100 mm tall, what are (*a*) the distance of the image from the lens and (*b*) the focal length of the lens? ●●

117. A converging lens for which $f = 50$ mm forms an image that is three times larger than the object. How far from the lens is the object? ●●

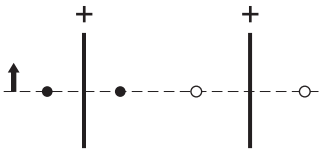
118. In vacuum, the wavelength of a certain light wave is 550 nm. (*a*) What is the wavelength in a medium where the wave speed is 2.4×10^8 m/s? What is the frequency of this light (*b*) in vacuum and (*c*) in the medium? ●●

119. The focal length of a lens can be different for different colors of light. (*a*) What optical principle causes this problem? (*b*) Can the problem be corrected? (*c*) Does the same problem affect mirrors? ●●

120. When an object is placed 1.5 m in front of a converging mirror, an image of the object appears twice as high behind the mirror. (*a*) Is the image real or virtual? (*b*) Is it upright or inverted? (*c*) What is the mirror's radius of curvature? ●●

121. The focal points of the two converging lenses shown in Figure P33.121 are denoted by solid dots for the left lens and open dots for the right lens. Draw a simplified ray diagram to locate the final image. ●●

Figure P33.121



122. Using a camera fitted with a lens for which $f = 50$ mm, you take a photograph of a person who is 1.7 m tall and is standing 29 m away from you. What is the height of the person's image on the film? ●●

123. A piece of electronic circuitry embedded in a layer of diamond ($n_{\text{diamond}} = 2.42$) emits light. The diamond is covered by a coating of flint glass ($n_{\text{flint}} = 1.65$). The interface between the two materials is flat, as is the interface between the flint glass and the surrounding air. The two interfaces are parallel. Relative to the direction normal to both interfaces, what are (*a*) the greatest angle at which light emitted in the diamond layer can enter the flint glass and (*b*) the greatest angle at which light emitted in the diamond layer travels through the flint glass and enters the surrounding air? ●●●

124. You and a friend are scuba diving on a calm day when you decide it would be neat to estimate the critical angle for light going from water into air. Your friend is underwater and has a laser pointer and a waterproof earpiece so that he can hear your instructions. He also has a floating marker attached to him that remains on the surface directly above him. Sitting in a small motorboat on the surface, you have a variety of tools for measuring distances (tape measures, fishing line that could be let out and later accurately measured, and so on). You think for a moment and decide on a course of action. ●●● CR

125. You are in charge of the projector in a movie theater. The film passes through a holder, and light from a bright bulb passes first through the film and then through a lens. You have several additional lenses that can be swapped out, and the position of the lens holder is easily adjustable, but moving the film holder seems more troublesome. A single film frame is 20.00 mm tall. You notice that when the lens is placed 100.0 mm from the film, a sharp image 1450 mm tall is formed on the movie screen on the other side of the lens. This image is far too small, filling less than one-third of the screen. If the patrons do not see a much larger, well-focused image, you may be returned to your previous job making popcorn. ●●● CR

126. For a lens in air, you have learned the conditions on its radii of curvature to make the lens converging or diverging. You have done some diving, and you realize that a lens has a different focal length underwater than it does in air. That starts you thinking about what would happen if a lens is placed in a medium with an index of refraction greater than that of the material of which the lens is made, but you are not even sure that it is possible to form images under such conditions. Lost in your thoughts, you drift off into dreamland. ●●● CR



34

Wave and Particle Optics

34.1 Diffraction of light
34.2 Diffraction gratings
34.3 X-ray diffraction
34.4 Matter waves
34.5 Photons

34.6 Multiple-slit interference
34.7 Thin-film interference
34.8 Diffraction at a single-slit barrier
34.9 Circular apertures and limits of resolution
34.10 Photon energy and momentum

CONCEPTS

QUANTITATIVE TOOLS

In Chapter 33, we considered the propagation of light along a straight path. The chapter title, "Ray Optics," reflects the fact that we considered propagating light only in the simplest way—as straight-line motion. You know from Chapter 30, however, that light is an electromagnetic wave. This means that it must undergo interference and diffraction, just like any mechanical wave. As you will learn in this chapter, light waves can interfere with one another and diffract when they pass through small openings.

Another fact about light you will learn in this chapter is that it has a dual nature: It is a wave, yes, but it also has the properties of a particle!

34.1 Diffraction of light

As we saw in Chapter 17, when a surface water wave strikes a barrier that has a small opening, the wave diffracts (spreads out) after it passes through the opening. **Figure 34.1*a***, for example, shows surface water waves diffracting nearly circularly after they pass through an opening.

Given that light is a wave, as we discussed in Chapter 30, why don't we see light diffract in a similar fashion after it travels through, say, a window? As Figure 34.1*b* shows, after passing through a window, light continues to travel in a straight line, casting a sharp-edged shadow with no discernible diffraction.

The reason light does not diffract through a window is that the wavelength of the light is very much smaller than the size of the window. In Figure 34.1*a*, the wavelength of the water wave is about the same as the width of the opening, but the wavelength of the light in Figure 34.1*b* is about a million times smaller than the width of the window.

Diffraction is indeed observed with light waves but only when the width of the opening through which the light passes is not much greater in size than the wavelength of the light. Empirical evidence shows that diffraction occurs through openings approximately two orders of magnitude greater than the wavelength. Thus, visible light, with a wavelength on the order of 1 μm, diffracts through apertures up to hundreds of micrometers wide.

To understand diffraction, it is useful to consider the propagation of wavefronts. As discussed in Sections 17.1 and 33.2, a wavefront is a surface on which a wave spreading through space has constant phase. Wavefronts are everywhere perpendicular to the direction of propagation of the wave. By convention, wavefronts are drawn at the crests of the waves, which means the separation between adjacent wavefronts equals the wavelength (Figure 17.2). Although in principle wavefronts can take any shape, usually we consider only those that are either planar or spherical. Most sources of light, from light bulbs to stars, can be modeled as point sources—single points that produce concentric spherical wavefronts. As discussed in Section 17.1 (see especially Figure 17.7), far away from a point source, the radius of the spherical wavefronts is so great that the wavefronts are very nearly planar. As a result, distant point sources can be considered to be sources of planar waves. Lasers also produce planar waves, even very close to the source. For this reason, we can use a laser beam in seeing how electromagnetic waves behave. Keep in mind, however, that our analysis applies to any type of electromagnetic waves, not just to laser beams.

Let us now determine under what conditions a planar wave spreads out as it propagates—in other words, under what conditions it undergoes diffraction. **Figure 34.2** shows planar wavefronts from a beam of electromagnetic waves propagating to the right, with point Q located at the upper end of the wavefronts and point P located outside the region reached by the wavefronts. Because a wavelet (see Section 17.4) centered on Q radiates toward P along the line QP, we expect the beam to spread out as it propagates. However, such spreading is not observed. The reason is that the wavelets centered on points below Q also radiate toward P, and we need to sum the

CONCEPTS

Figure 34.1 Surface water waves diffract when they pass through a gap, whereas light coming in through a window seems not to diffract—it forms a sharp-edged shadow. Notice that the gap in the barrier is roughly as wide as the wavelength of the water waves.

(*a*) (*b*)

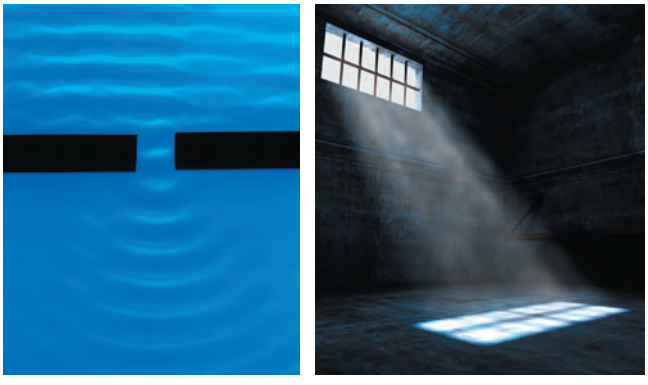
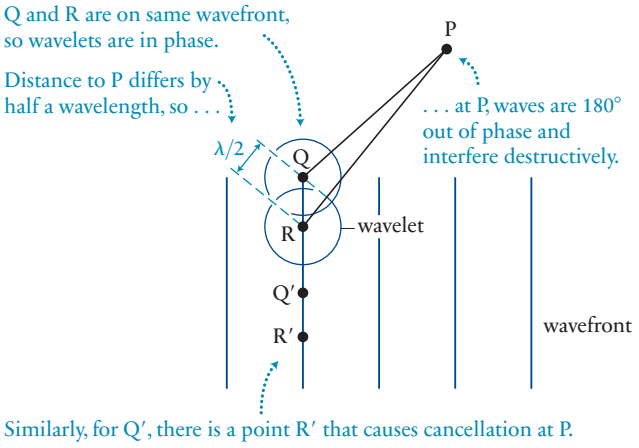


Figure 34.2 The reason we don't usually see diffraction for light beams, provided the beam is very much wider than the wavelength of the light waves.



contributions of all these wavelets. Consider, for example, the wavelet centered on R, for which the distance PR is exactly half a wavelength longer than the distance PQ. Because Q and R lie on the same wavefront, they produce coherent wavelets. (You should review Section 17.4 if you do not see why this is true.) This means that at P the electric field part of the wavelet from Q is 180° out of phase with the electric field part of the wavelet from R. Thus the two electric fields interfere destructively at P. Because the same is true for the magnetic field part of the wavelets, there is complete destructive interference (see Section 16.3) at P.

In the same manner, for points Q′ lying below Q on the wavefront, we can find on the same wavefront a point R′ for which the distance R′P is exactly half a wavelength longer than the distance Q′P. Thus the fields from the wavelet traveling along R′P cancel those from the wavelet traveling along Q′P. If the wavefronts extend far enough below Q, we can always find points that cancel the waves from any other point. As a result, we conclude that the light does not spread outside the beam; in other words, there is no diffraction.

34.1 Consider the point P located ahead of the wavefronts shown in **Figure 34.3**. Following the line of reasoning used in the preceding discussion, what can you say about the intensity (power/area, as defined in Eq. 30.34 and Section 17.5) at P once the wave fronts reach it? (Hint: Consider separately points R above, below, and on the ray through P.)

Checkpoint 34.1 demonstrates that the intensity of a planar wave is uniform as the wave propagates forward because the fields from individual wavelets cancel in any outward direction and reinforce only in the direction of propagation. Section 17.3 (especially Figure 17.20) shows that combining the waves from many adjacent point sources indeed produces a planar wave that propagates forward with very little spreading. This is true because the wavelets interfere destructively with one another in any direction other than the direction of propagation of the wavefronts.

The cancellation process we used in Figure 34.2 can be used only when the width of the laser beam, and hence the width of the wavefronts, is much greater than the wavelength of the light. That way, all points on a wavefront can be paired with other points on that wavefront that are half a wavelength or more away, and the wavelets from these points cancel.

Figure 34.3 Checkpoint 34.1.

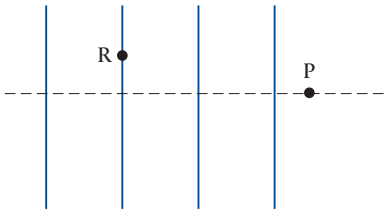
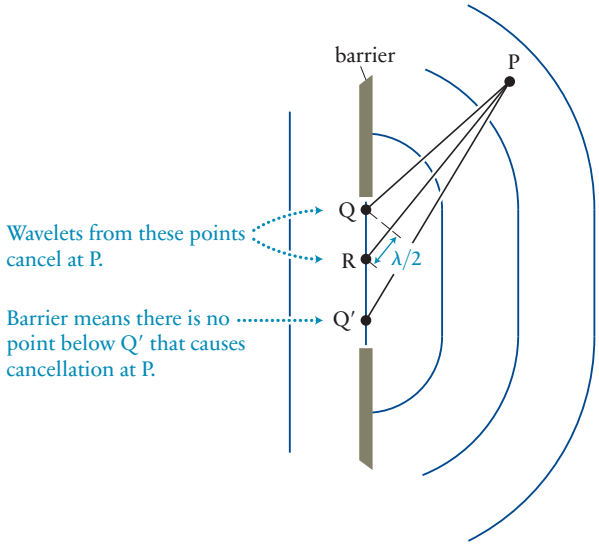


Figure 34.4 The reason we *do* see diffraction when a light beam is transmitted through a small aperture.



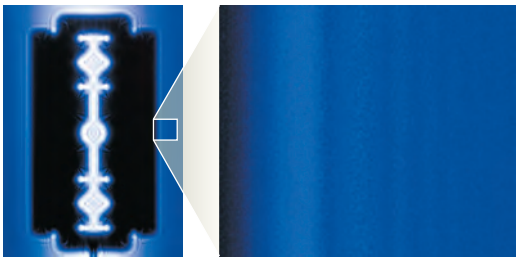
If the width of the beam is comparable to the wavelength of the light, not all points on the wavefronts can be paired in this manner. To see why this is so, let us place a barrier in front of our light source, as in **Figure 34.4**. You can think of this drawing as a bird's-eye view of a beam of light traveling to the right and running into a wall that has a gap in it. Light hitting the wall on either side of the gap cannot pass through. As each wavefront of the beam reaches the barrier, only the portion that hits the gap continues moving to the right. The width of each wavefront that passes through the gap is equal to the gap width.

Suppose the gap in Figure 34.4 has a width equal to 2λ. All waves that reach P from wavelets that originate above Q′ cancel, as indicated by the rays drawn from points Q and R in Figure 34.4; wavelets that originate at or below Q′ are not canceled at P because those wavelets lack corresponding wavelets at an appropriate distance below Q′. As a result, some of the light spreads out past the edges of the original path of the beam—the light is diffracted. Diffraction in this situation occurs for exactly the same reason as the diffraction of water waves pictured in Figure 34.1*a*. Diffraction of light occurs when a planar wave passes through an aperture that is only micrometers wide (much less than the width of a hair). As shown in Figure 34.1*a*, if the width of the aperture is equal to or less than the wavelength, the wavefronts coming from the aperture are spherical. If the aperture is a few wavelengths wide, the wavefronts are elongated right after they pass through the aperture, as shown in Figure 34.4.

An ordinary window, like that shown in Figure 34.1*b*, is effectively infinitely wide relative to the wavelength of the light, so only the light at the very edges of the window diffracts. In practice, not even diffraction from the edge of the window is observed because the edge is not perfectly smooth on a micrometer scale. However, it *is* possible to observe diffraction of light from the edge of a smooth razor blade, as shown in **Figure 34.5** on the next page.

CONCEPTS

Figure 34.5 Edge diffraction is not usually apparent for visible light because most edges are not smooth enough. However, it can be observed around the edge of a razor blade. The blade in this image is illuminated by a point source of monochromatic light.



EXAMPLE 34.1 Spreading out

Do you expect to be able to observe the diffraction of light through (*a*) the front door to your house; (*b*) the holes in a button; (*c*) the gaps between threads of the fabric of an umbrella?

❶ **GETTING STARTED** I expect to see noticeable diffraction through openings up to roughly two orders of magnitude times the wavelength of the light. I therefore need to estimate the width of each opening and determine the width-to-wavelength ratio.

❷ **DEVISE PLAN** To estimate the width of the front door and of the holes in the button, I can draw on my experience; to estimate the widths of the gaps between the threads of the fabric of an umbrella, I shall use an upper limit. Then I shall take the ratios of these widths to the wavelength of light in the middle of the visible range, 500 nm.

❸ **EXECUTE PLAN** (*a*) A door is about 1 m wide, so the ratio of the door's width to 500 nm is 2×10^6, much too great to see diffraction. ✔

(*b*) The holes in a button are about 1 mm in diameter, so the ratio of this width to 500 nm is 2×10^3, still too great to see diffraction. ✔

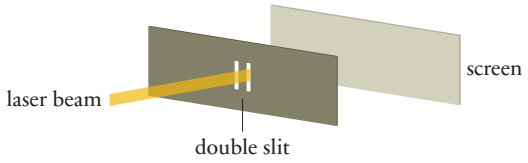
(*c*) I know that a human hair, which is less than 100 μm in diameter, cannot easily be threaded through a piece of fabric. Therefore I estimate the gaps between the threads in the fabric to be one-tenth of a hair diameter, or 10 μm at most. The ratio of a gap width to 500 nm is therefore 20 or less, and I expect to see diffraction through the gaps in the fabric. ✔

❹ **EVALUATE RESULT** I know from experience that I do not see diffraction through a doorway. I can check my answers to parts *b* and *c* by looking through the holes in a button and through an open umbrella. When I do so, I see no diffraction through the button but I can see diffraction through the umbrella if the fabric is dark. (This diffraction is particularly noticeable when I look at a streetlight at night through the umbrella.)

CONCEPTS

34.2 In discussing how a planar wave propagates, we could turn our earlier argument around and say that for each point Q in Figure 34.2 there is a point S somewhere on the wavefront that radiates toward P along a path exactly one wavelength longer than that from Q, and therefore there should be a nonzero intensity at P. What is wrong with this argument?

Figure 34.6 When the planar electromagnetic waves of a laser beam pass through a pair of narrow slits, what do we see on the screen?



34.2 Diffraction gratings

What happens if instead of passing through a single small aperture, a planar electromagnetic wave strikes a barrier that contains two narrow slits at normal incidence, as shown in **Figure 34.6**? If the slit width is much less than the wavelength of the wave, the slits serve as two coherent point sources of electromagnetic waves of the same wavelength as the wave striking the barrier, and the waves from the two sources interfere with one another in the manner described in Section 17.3 for two adjacent sources of surface waves. The only difference is that electromagnetic waves are not confined to a planar surface but spread out in three dimensions. If the two slits are either round or square, the waves that emerge from them are spherical. If the slits are much taller than they are wide, as in Figure 34.6, the slits serve as lines of point sources and the waves that emerge from the slits have cylindrical wavefronts.

The crests of the waves from these two coherent point sources overlap in certain directions, as shown in **Figure 34.7** (and Figure 17.13). Between these directions where there is overlap are directions along which the waves from the two sources cancel. If we place a screen to the right of the slits, a pattern of alternating bright and dark bands appears on the screen, as shown in **Figure 34.8**. Such dark and bright bands are commonly called **interference fringes.** The bright fringes are labeled by a number called the **fringe order** m. The central bright fringe is the *zeroth-order bright fringe* ($m = 0$); around it are higher-order bright fringes ($m = 1, 2, \ldots$). Note that the pattern is symmetrical about this zeroth-order bright fringe.

Figure 34.7 Interference between the diffracted waves emerging from the two slits.

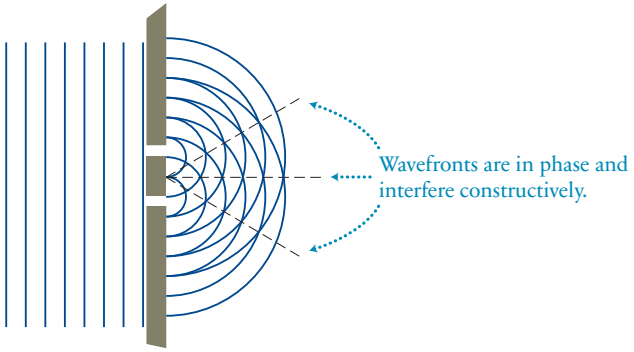
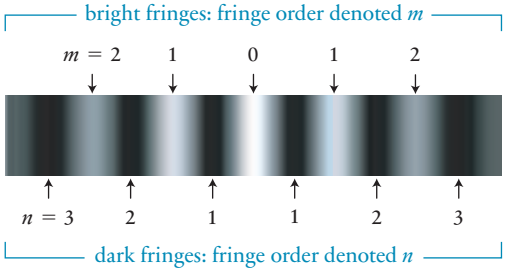


Figure 34.8 The interference pattern produced when the laser beam of Figure 34.6 passes through a pair of slits and strikes the screen.



How do we determine the locations of the bright fringes? The central (zeroth-order, $m = 0$) bright fringe is simplest to locate. As shown in Figure 34.7, waves from the two sources interfere constructively along a perpendicular line running through the midpoint between the two slits. Because the waves coming from the two slits travel the same distance to reach any point along this perpendicular line, the waves arrive in phase with each other.

To locate the other bright interference fringes, we need to work out the directions in which the difference in path length between waves coming from the upper source and waves coming from the lower source is an integer number of wavelengths. In these directions, constructive interference between waves from the two sources produces bright fringes.

Let's consider two rays, one from each slit, that meet at the screen to form a fringe, as shown in **Figure 34.9a**. In general we shall take the distance from the sources to the screen to be much greater than the distance between the slits. Note from the figure how, even though the rays eventually meet at the screen, their paths are essentially parallel when they emerge from the slits (Figure 34.9*b*). If we denote the angle between the nearly parallel rays and the normal to the barrier as θ, we can say that the difference in path length for waves emitted at angle θ from the two sources is $d \sin \theta$, where d is the distance between the centers of the slits (Figure 34.9*b*). When this path-length difference is equal to an integer multiple of the wavelength, constructive interference occurs. The central bright fringe corresponds to $d \sin \theta = 0$ or $\theta = 0$. Subsequent bright fringes are located at angles given by $d \sin \theta_m = \pm m\lambda$, where $m = 1, 2, \ldots$ denotes the order of the bright fringe. The plus and minus signs give rise to bright fringes on either side of the central maximum.

Likewise, in directions that correspond to path-length differences of an odd number of half-wavelengths, waves from the two sources interfere destructively, resulting in dark fringes. For these fringes, we use n to denote fringe order. The smallest angle at which destructive interference occurs corresponds to $d \sin \theta = \frac{1}{2}\lambda$. More generally, the angles at which dark fringes occur are given by $d \sin \theta_n = \pm (n - \frac{1}{2})\lambda$, where $n = 1, 2, \ldots$ denotes the order of the dark fringe. The dark fringes around the zeroth-order bright fringe are the first-order dark fringes.

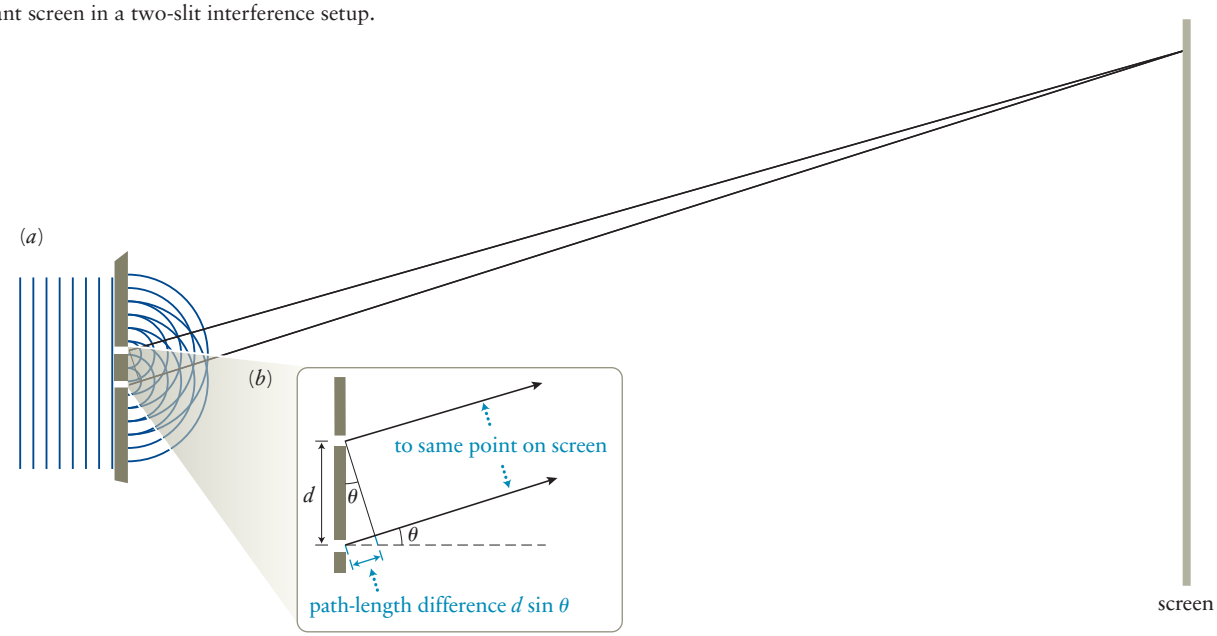
EXAMPLE 34.2 Two-slit diffraction grating

Coherent green light of wavelength 530 nm passes through two very narrow slits separated by 1.00 μm. What is (*a*) the angular location of the first-order bright fringe? (*b*) the angular separation between the $n = 1$ and $n = 2$ dark fringes?

❶ **GETTING STARTED** This problem involves interference between light rays passing through two closely spaced, very narrow slits, as shown in Figure 34.9.

(Continued)

Figure 34.9 Determining the path-length difference between two rays traveling to a point on a distant screen in a two-slit interference setup.



CONCEPTS

2 DEVISE PLAN The angular locations of the centers of the bright fringes are given by the condition for constructive interference, $d \sin \theta_m = m\lambda$; the angular locations of the dark fringes are given by the condition for destructive interference, $d \sin \theta_n = (n - \frac{1}{2})\lambda$. (For both bright and dark fringes, I omit the $\pm$ signs because I'll only consider the fringes on one side of the central maximum.) I therefore need to use these relationships to calculate the angular locations of the fringes. To use the constructive and destructive interference conditions, I need to identify the appropriate values of m and n and then use them with the wavelength and the distance between slits to obtain the angular positions of the bright and dark fringes I am interested in.

3 EXECUTE PLAN (*a*) The first-order bright fringe corresponds to $m = 1$, which means the center of the first-order bright fringe is located at the value θ_1 corresponding to $d \sin \theta_1 = \lambda$. Substituting for d and λ in this expression and solving for θ_1, I obtain

$$\theta_1 = \sin^{-1}\left(\frac{0.530\ \mu\text{m}}{1.00\ \mu\text{m}}\right) = 32.0°. \checkmark$$

(*b*) The two lowest-order dark fringes correspond to $n = 1$ and $n = 2$, and their centers occur at the angles corresponding to $d \sin \theta_1 = \lambda/2$ and $d \sin \theta_2 = 3\lambda/2$. Substituting and solving, I obtain

$$\theta_1 = \sin^{-1}\left(\frac{0.530\ \mu\text{m}}{2 \times 1.00\ \mu\text{m}}\right) = 15.4°$$

$$\theta_2 = \sin^{-1}\left(\frac{3 \times 0.530\ \mu\text{m}}{2 \times 1.00\ \mu\text{m}}\right) = 52.7°.$$

The angular separation is thus $52.7° - 15.4° = 37.3°$. ✔

4 EVALUATE RESULT The center of the $m = 1$ bright fringe is roughly halfway between the centers of the $n = 1$ and $n = 2$ dark fringes, as I expect from Figure 34.8.

CONCEPTS

34.3 Does the spacing of the bright fringes in the two-slit arrangement in Figure 34.6 increase, decrease, or stay the same if we (*a*) increase the spacing d of the slits, or (*b*) increase the wavelength λ of the light incident on the arrangement?

Now consider the effect of many equally spaced narrow slits in a barrier on which planar waves are incident normally (**Figure 34.10**). Once again, we can determine

Figure 34.10 Path-length difference for planar waves striking a barrier with multiple slits.

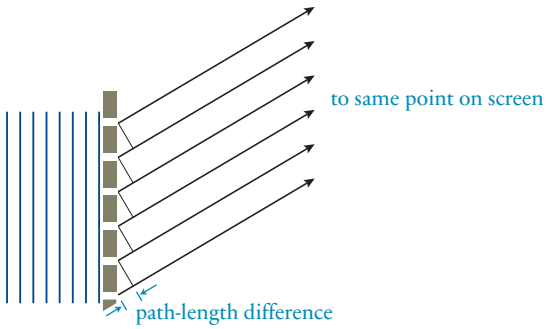
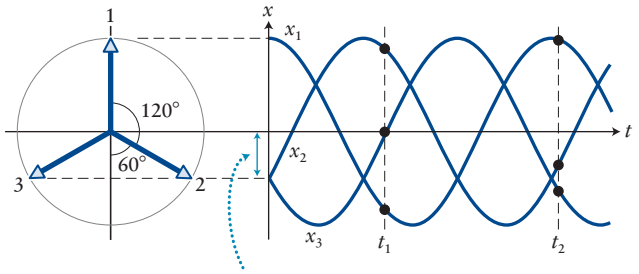


Figure 34.11 Three coherent waves interfere destructively when each is out of phase with the other two by one-third of a cycle.



Projection of phasor 2 onto vertical = cos 60° times amplitude = half amplitude.

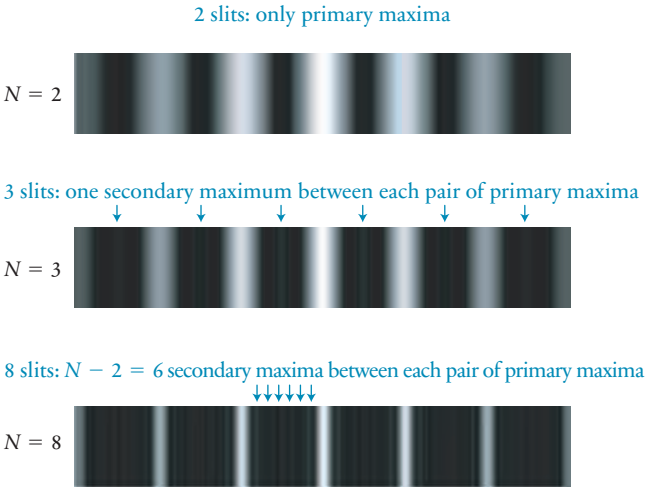
the fringe pattern produced at a given location on a distant screen by combining all the waves (one from each slit) that travel to that location. The condition for constructive interference among all the waves is equivalent to that for two slits. As can be seen from Figure 34.10, the path-length difference between each pair of adjacent waves is the same. This means that if the waves from two of the slits interfere constructively, the waves from *all* of the slits do the same. Bright fringes therefore appear at angles corresponding to $d \sin \theta_m = \pm m\lambda$, with d the separation between adjacent slits. The location of bright fringes therefore does not depend on the number of slits as long as the separation d between adjacent slits is the same for all slits.

34.4 Suppose there are three slits in a barrier on which light is incident normally, with each slit separated from its neighbor by a distance d. Do the waves from all three slits cancel perfectly at angles given by $d \sin \theta = \pm (n - \frac{1}{2})\lambda$?

As Checkpoint 34.4 illustrates, the condition for complete destructive interference of all of the waves is not the same for three slits as for two. Instead, as **Figure 34.11** shows, three coherent waves cancel one another perfectly when each is out of phase with the other two by one-third of a cycle. The three waves also cancel one another perfectly when each is out of phase with the other two by two-thirds of a cycle. In that case, phasors 2 and 3 in Figure 34.11 are interchanged. So there are two dark fringes between each pair of bright fringes. In between these two dark fringes is a faint bright fringe (**Figure 34.12**). The brightest fringes are the *principal maxima* in the interference pattern. These correspond to constructive interference of the waves diffracted by all three slits. The fainter bright fringes are *secondary maxima*. At these locations the cancellation is not complete.

Four coherent waves cancel one another when adjacent waves are out of phase by one-fourth of a cycle. In the same manner, if there are N slits, the condition for complete destructive interference is that each wave must differ in phase by $1/N$ of a cycle from its immediate neighbors. Then the N waves are evenly distributed

Figure 34.12 Interference pattern caused by the diffraction of a coherent beam of light through two, three, and eight narrow slits.



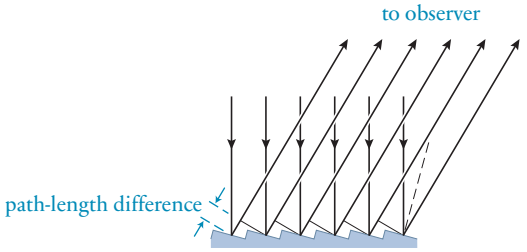
throughout one cycle of oscillation and add to zero. The condition for the path-length differences for the dark fringes is thus $d \sin \theta_k = \pm (k/N)\lambda$, where k is any integer that is *not* a whole-number multiple of N (because when $k/N = m$, we have constructive interference).

Although the bright fringes are in the same location regardless of the number of slits, there are now $N - 1$ dark fringes between the bright fringes and $N - 2$ secondary maxima between each pair of principal maxima. As a result, as N increases, the bright fringes become narrower and brighter, as shown in Figure 34.12. (The brightness of the pattern corresponds to the intensity of the light striking the screen.)

34.5 Why does the brightness of the fringes increase as the number of slits increases?

The interference of a planar electromagnetic wave as it passes through many closely spaced narrow slits is due to the diffraction that occurs at the slits. A barrier that contains a very large number of such slits is therefore called a **diffraction grating.** Diffraction gratings can be either transmissive (such as the one shown in Figure 34.10) or reflective. Reflective diffraction gratings are made by engraving grooves to reflect light from a surface, as shown in **Figure 34.13**. The grooves

Figure 34.13 Reflective diffraction grating.



on a digital video disc (DVD) form a reflective diffraction grating.

Why is the light reflected from the compact disc surface so colorful? You found in Checkpoint 34.3 that the position of interference fringes produced by light of a single color depends on the wavelength. When white light, which contains many different colors and therefore many different wavelengths, falls on a diffraction grating, the fringes for each wavelength are displaced from each other, producing a series of rainbows.

34.6 Suppose the light striking the reflective diffraction grating in Figure 34.13 is white light—that is, light consisting of all the colors of the rainbow. Red light has the longest wavelength of the colors that make up white light; violet has the shortest. (*a*) Is the angle at which first-order constructive interference occurs for violet light less than, equal to, or greater than the angle for red light? (*b*) Why are multiple rainbows visible in the reflected light?

Very precisely manufactured diffraction gratings have many uses in scientific equipment. They are most widely used to disperse visible, ultraviolet, or infrared light into its constituent wavelengths because the resulting spectrum provides information about the object that emitted the light. For example, astronomers often use a diffraction grating attached to a telescope to identify the wavelengths present in the light from stars, in order to understand the chemical composition of the stars or their distance from Earth.

Certain common objects can also function as diffraction gratings. For example, if you look through a piece of dark, finely woven, taut fabric at a point source of light (say, a distant street light through the fabric of a dark umbrella), you will see fringes.

34.7 Diffraction gratings used in astronomical instruments must be able to separate wavelengths that are quite close together. (*a*) To increase the ability to do this, should the separation between slits be made less or greater? (*b*) Does the width of the slits affect the diffraction pattern?

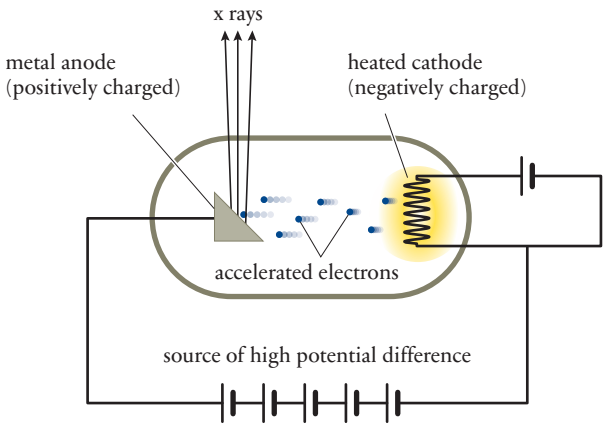
CONCEPTS

34.3 X-ray diffraction

The interference of very-short-wavelength electromagnetic radiation is widely used to study the structure of materials. **X rays** are electromagnetic waves that have wavelengths ranging from 0.01 nm to 10 nm, more than 100 times less than the wavelengths of visible light.

Figure 34.14 on the next page shows one way to generate x rays. Electrons are ejected from a heated cathode on the right and accelerated by a potential difference of several thousand volts toward a metal anode on the left. The electrons crash into the atoms

Figure 34.14 Schematic diagram for a cathode ray tube x-ray emitter.



that make up the anode; this inelastic collision decelerates the electrons very rapidly and gives the atoms a great deal of internal energy. The atoms then re-emit this energy in the form of x rays. In addition, the rapidly decelerating electrons radiate x rays, typically at a 90° angle to the path of the accelerated electrons.

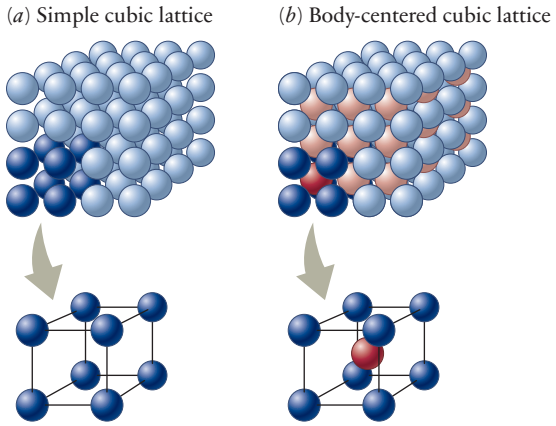
X rays can pass through many materials with low mass density that are opaque to visible light. For example, they pass through soft tissues in the human body but are strongly absorbed by bones and teeth. As a result, x rays are widely used to obtain photographic images of the skeleton (**Figure 34.15**). X-ray imaging of blood vessels or soft internal organs can be done by giving the patient a drug containing dense elements, such as iodine, because dense material absorbs x rays. For example, one way cardiologists diagnose heart problems is to directly observe blood vessels in a patient's heart by injecting a dense-atom drug into the patient's blood and taking x-ray movies of the beating heart as the blood is pumped through.

Because x-ray wavelengths are either shorter than or comparable to the typical distance between atoms in solid materials (0.1 nm to 1 nm), x-ray diffraction can be used to study atomic arrangements in solids. Many solids are

Figure 34.15 Bones and teeth absorb x rays, whereas soft tissues are nearly transparent to them.



Figure 34.16 Two examples of crystal lattices.



crystalline, meaning that their atoms are arranged in a three-dimensional, regularly spaced grid called a *crystal lattice* (**Figure 34.16**). The lattice serves as a three-dimensional diffraction grating for x rays because the lattice spacing is comparable to the x-ray wavelength.

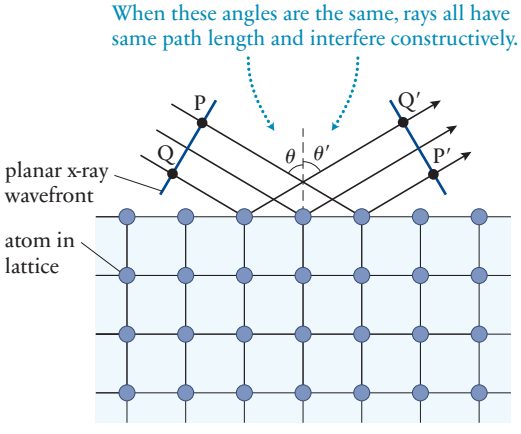
Consider what happens when x rays strike the top plane of atoms in a crystal lattice* at an angle θ (**Figure 34.17**). Each atom that is struck by the beam of x rays acts as the source of a wavelet emitting waves in all directions, much like the slits of the diffraction gratings we discussed in the preceding section. Waves emitted at $\theta' = \theta$ have the same path length and so they add constructively, yielding a strong reflected beam.

34.8 Considering only the top row of atoms in Figure 34.17, are there any other directions in which the x rays diffracted by the atoms interfere constructively?

As you saw in Checkpoint 34.8, x rays diffracted by the crystal in directions other than at angle $\theta' = \theta$ are much weaker than those diffracted at angle $\theta' = \theta$. We

*The lattice shown is a so-called cubic lattice. Lattices can be much more complex than the cubic lattice, but the principles of diffraction are the same.

Figure 34.17 Diffraction of x rays by the atoms at the surface of a crystal lattice.



CONCEPTS

Figure 34.18 Interference of x rays diffracted by adjacent planes of a crystal.

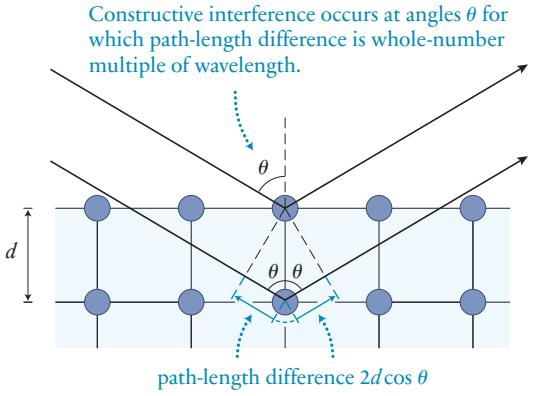
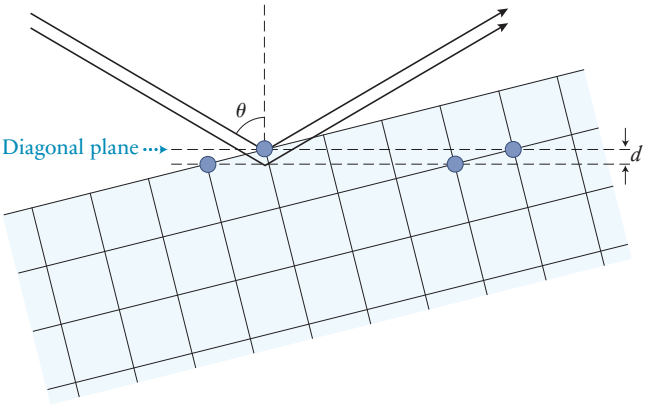


Figure 34.19 Constructive interference of x rays diffracted by two diagonal crystal planes.



might therefore expect not to see any x-ray diffraction from crystals at other angles. However, a crystal consists of many planes of atoms, and some incident x rays penetrate into the crystal. We thus need to take into account the diffraction of the x rays by the atoms in multiple crystal planes to determine the diffraction of the crystal as a whole.

Figure 34.18 shows the x rays diffracted by atoms in two adjacent planes of a crystal. For most angles of incidence, the waves diffracted by atoms in different planes differ in phase and so interfere destructively. However, when the difference in path length between rays diffracted by atoms in different planes is a whole-number multiple of the x rays' wavelength, the rays are in phase and interfere constructively. The path-length difference equals $2d \cos \theta$, where d is the distance between adjacent planes. Therefore the condition for constructive interference is $2d \cos \theta = m\lambda$. This condition is called the **Bragg condition,** after the father-and-son team of physicists who formulated it. Because the atomic spacing and the x-ray wavelength are fixed, crystals reflect x rays only at those angles for which $2d \cos \theta$ is an integer multiple of the x-ray wavelength.

The atoms in the lattice of a crystal define many different lattice planes. **Figure 34.19** shows two sets of lattice planes in a cubic crystal, one indicated by dashed lines (planes parallel to the surface of the crystal) and the other indicated by solid lines (diagonal planes). If we tilt the crystal so that the angle the incident x rays make with its surface is different from the angle shown in Figure 34.18, a different set of planes with a different spacing can produce constructive interference of the diffracted waves.

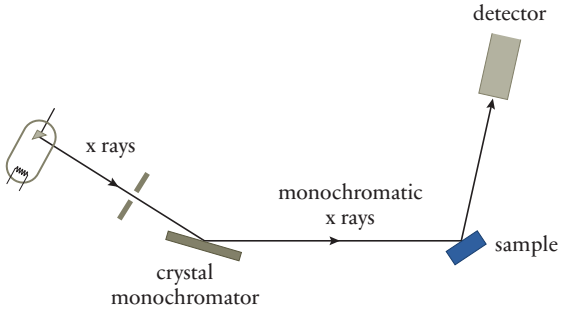
By measuring the angles at which strong x-ray diffraction occurs, one can determine the arrangement of atoms in a crystalline solid. **Figure 34.20** shows how such a measurement is carried out. An x-ray source like that shown in Figure 34.14 is used to produce a beam of x rays of various wavelengths. The beam is then diffracted from a *crystal monochromator* (which is simply a crystal of known lattice spacing) positioned at an angle chosen so that the Bragg condition is satisfied for one desired wavelength of x rays. Because the other wavelengths in the original beam do not satisfy the Bragg condition, a *monochromatic* (single-wavelength) beam of x rays is diffracted from the monochromator to the sample.

The sample of crystalline material whose lattice is being studied is slowly rotated with respect to the monochromatic beam, and as this rotation takes place, the intensity of x rays diffracted from the sample is measured on a detector as a function of the angle α between the x rays and the sample surface (Figure 34.20*b*). This angle is often called the *Bragg angle* α. As this angle changes, the Bragg condition is generally not satisfied. However, at specific Bragg angles, the various crystal planes in the

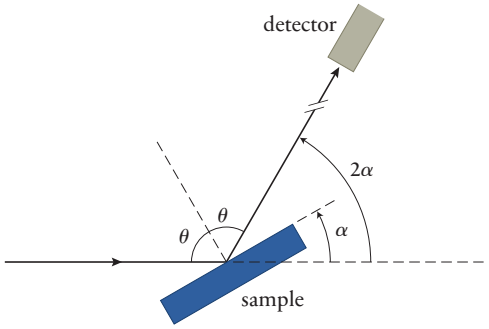
CONCEPTS

Figure 34.20 (*a*) Apparatus for studying x-ray diffraction from a crystalline solid. (*b*) Relationship between the incident angle θ and the Bragg angle α.

(*a*) Apparatus for studying x-ray diffraction from a crystalline solid



(*b*) Relationship between incident angle θ and Bragg angle α



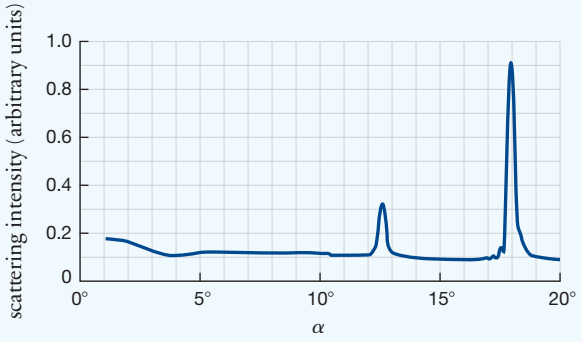
sample satisfy the Bragg condition, producing a high intensity of x rays on the detector.

34.9 Express the Bragg condition $2d \cos\theta = m\lambda$ in terms of the Bragg angle α between the x rays and the surface of the sample rather than the angle θ between the x rays and the surface normal.

EXAMPLE 34.3 X-ray diffraction

Figure 34.21 shows diffracted x-ray intensity as a function of the Bragg angle α, obtained using x rays having a wavelength of 0.11 nm. (*a*) Without calculating values for the lattice spacing d, identify which of the two peaks corresponds to a greater distance between adjacent planes in the sample being studied. (*b*) Calculate the distance between adjacent planes corresponding to each peak.

Figure 34.21 Example 34.3.



❶ **GETTING STARTED** The peaks in a graph of x ray intensity as a function of the Bragg angle α correspond to constructive interference and therefore values of α that satisfy the Bragg condition.

❷ **DEVISE PLAN** For part *a*, I can use the Bragg condition, together with the shape of the graph, to deduce which peak results from the greater d value. Because the graph gives intensity in terms of the Bragg angle, I shall need the Bragg condition expressed in terms of the Bragg angle.

For part *b*, I can solve this form of the Bragg condition for d and then insert my two given α values to determine the plane separation distance in each case. Looking at this form of the Bragg condition, I see that the Bragg angle α at which a peak occurs increases as m increases. Because this graph begins at $\theta = 0°$, the two peaks must correspond to $m = 1$ for the interference patterns when the crystal surface is oriented at the two Bragg angles I am working with.

❸ **EXECUTE PLAN** (*a*) From Checkpoint 34.9 I know that the Bragg condition in terms of the Bragg angle is $2d \sin\alpha = m\lambda$. In order for the product $2d \sin\alpha$ to remain constant, α must decrease as the distance d between adjacent planes increases. Therefore the peak at the smaller α value corresponds to the greater distance between planes. ✔

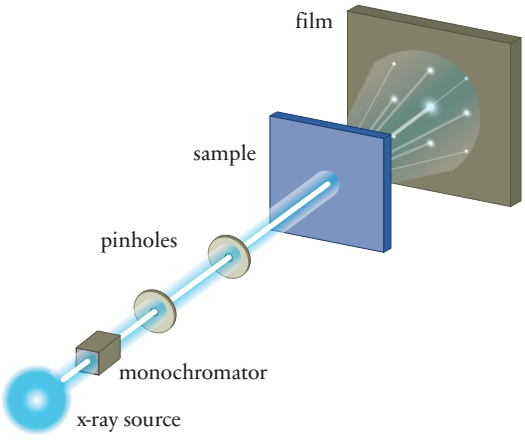
(*b*) To obtain d for each peak, I solve $2d \sin\alpha = m\lambda$ for d and then substitute $m = 1$ and the values for α and λ. For the short peak, $\alpha = 12.5°$, which gives $d = \lambda/(2 \sin\alpha) = (0.11 \text{ nm})/(2 \sin 12.5°) = 0.25 \text{ nm}$. For the tall peak, $\alpha = 18°$, so $d = \lambda/2 \sin\alpha = (0.11 \text{ nm})/(2 \sin 18°) = 0.18 \text{ nm}$. ✔

❹ **EVALUATE RESULT** The Bragg angle α at which constructive interference occurs decreases as the distance between planes increases. This is consistent with what I found previously for interference between two slits (Checkpoint 34.3), in which increasing the distance between slits also causes the angle between fringes to decrease. In general, the size of an interference pattern decreases as the distances between interfering sources increase. Finally, the smaller value of d multiplied by $\sqrt{2}$ gives the greater value of d, as it should for the distances between planes for the cubic lattice in Figure 34.19.

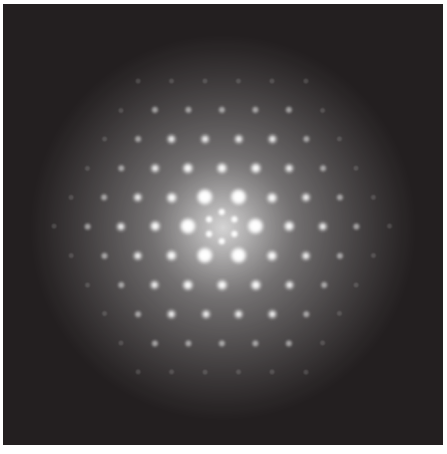
Many studies of crystal structure are done by passing x rays through the crystal rather than reflecting them from the various crystal planes. The crystal then acts like a three-dimensional transmissive diffraction grating for the x-ray beam. Instead of a single line of slits, the beam encounters many rows of slits. As a result, many rows of fringes usually called "spots" are formed. The experimental apparatus for such a measurement is shown in **Figure 34.22a**.

Figure 34.22 X-ray crystallography.

(*a*) Schematic apparatus for x-ray crystallography



(*b*) X-ray diffraction pattern of diamond lattice



CONCEPTS

From the position and intensity of the spots in the resulting diffraction pattern (Figure 34.22*b*), we can deduce the arrangement of the atoms in the sample.

The earliest x-ray diffraction studies of crystals were done on simple crystals, such as metals that form cubic or other very simple lattices. However, x-ray diffraction has also been used to study the atomic structure of much more complicated molecules, by crystallizing solutions of such molecules. The double helical structure of DNA was determined using x-ray diffraction, also called *x-ray crystallography*, and crystallography is widely used today to determine the atomic structure of even more complicated biological molecules.

34.10 While comparing the x-ray diffraction patterns from two crystals, you determine that crystal A produces a pattern with more widely spaced diffraction spots than crystal B. Which crystal has the greater atomic spacing?

34.4 Matter waves

Patterns remarkably similar to x-ray diffraction patterns can be obtained by aiming a narrow beam of electrons at a crystal. **Figure 34.23** shows the pattern obtained by sending a beam of high-speed electrons through a solid crystalline sample in an instrument called an electron microscope. The shape of the pattern is similar to that obtained with x rays (Figure 34.22*b*), which tells us that electrons are also diffracted by crystals. This discovery, in turn, suggests that the electrons behave like waves because interference and diffraction are wave phenomena. Indeed, electrons have been found to exhibit interference in many other experiments. For example, a beam of electrons aimed at two very narrow slits produces an interference pattern similar to that of a beam of light aimed at two narrow slits (**Figure 34.24**).

Varying the speed of the electrons changes the spacing of the interference pattern, which indicates that the electron wavelength depends on speed.

Figure 34.23 Electron diffraction pattern for a diamond lattice. Notice the similarity to the x-ray pattern in Figure 34.22*b*.

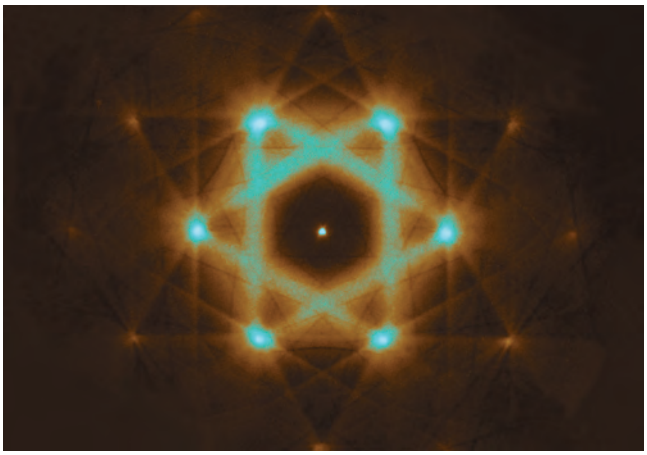
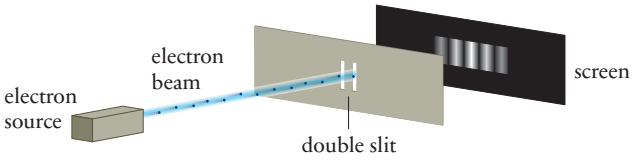


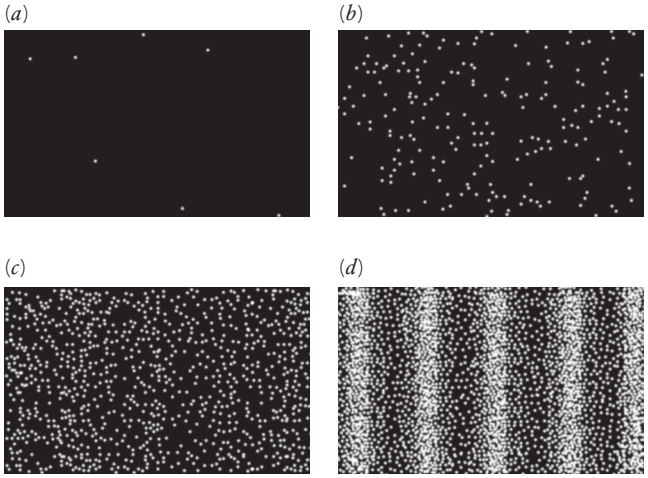
Figure 34.24 Apparatus for observing two-slit interference with an electron beam.



34.11 Spots in an electron diffraction pattern, such as that in Figure 34.23, move closer together as the speed of the electrons is increased. Does this mean the wavelength of the electrons increases or decreases with increasing speed?

Up to now we have considered electrons as being nearly pointlike particles, so it is surprising—to say the least—to discover that they also behave like waves. In fact, electrons can exhibit both particle behavior and wave behavior in a single experiment! This dual behavior was vividly demonstrated in an experiment done in 1989 in Japan using an apparatus similar to that shown in Figure 34.24. The number of electrons emitted by the source was kept very low, so as to ensure that at any given instant at most one electron was traveling from the source to the screen. The screen shows the place where each electron arrives as a bright dot, and at first these dots appear at seemingly random locations, as shown in **Figures 34.25*a*** and *b*. However, as more and more electrons reach the screen one after another, it becomes clear that the electron impacts are not randomly located (Figures 34.25*c* and *d*). Rather, they are arranged exactly in the two-slit interference pattern observed for a higher-intensity electron beam. Covering one slit and forcing the electrons to pass through the other makes the interference pattern disappear, just as it would for a light wave or a water wave.

Figure 34.25 When we perform two-slit diffraction with a very weak electron beam, we can see the pattern build up over time. At first (*a*, *b*), the dots that mark electron impacts seem to be scattered randomly, but as more accumulate (*c*, *d*), the diffraction pattern becomes evident.



CONCEPTS

Experiments show that after many electrons are allowed to pass one at a time through the apparatus, the number of electrons arriving on each small region of the screen per unit time is proportional to the intensity that would be observed at that region if light of the appropriate wavelength were shone on the slits. In other words, the *probability* of any single electron arriving at that small region in a fixed time interval corresponds to the intensity of the two-slit interference pattern at that location.

The observation of individual electrons arriving at the screen one after another indicates that the electrons are individual particles. However, if they are particles in the sense we think of for material objects, they must pass through either the right slit or the left slit, and so how can they produce an interference pattern?

The very counterintuitive conclusion is that each electron somehow travels through *both* slits simultaneously! Such a statement is not surprising for a wavefront hitting the two slits, but it goes completely against our intuitive notion of what we call a "particle."

We cannot explain the results of this experiment by concluding that electrons are *waves* because a classical wave could not produce the individual pinpoint images on the screen shown in, say, Figure 34.25*a*. If electrons were waves, the full diffraction pattern would be visible from the start (although it would be very faint). Scientists have concluded that electron behavior can be explained only if an electron has both particle properties and wave properties. This **wave-particle duality**—the possession of both wave properties and particle properties—has been observed not only for electrons but also for all other subatomic particles, for individual atoms, and, as we shall see in the next section, for light.

An expression for the wavelength of a particle, called the **de Broglie wavelength,** was proposed in 1924 by the French physicist Louis de Broglie* and confirmed experimentally a few years later. The de Broglie wavelength is inversely proportional to the momentum of the particle, $\lambda = h/p$, and the proportionality constant, called **Planck's constant,** is $h = 6.626 \times 10^{-34}\ \text{J}\cdot\text{s}$. That this constant is a very small number indicates that the wavelength of macroscopic objects is extremely small. Because waves exhibit diffraction and interference only on length scales comparable to their wavelength, the wave nature of matter has been observed only with subatomic particles, atoms, and molecules. Exercise 34.4 illustrates this point.

EXERCISE 34.4 Electron versus baseball

Calculate the de Broglie wavelength associated with (*a*) a 0.14-kg baseball thrown at 20 m/s and (*b*) an electron of mass 9.1×10^{-31} kg moving at 5.0×10^{6} m/s.

SOLUTION I use the expression $\lambda = h/mv$ for the de Broglie wavelength.

(*a*) $$\lambda_{\text{de Broglie, baseball}} = \frac{6.626 \times 10^{-34}\ \text{J}\cdot\text{s}}{(0.14\ \text{kg})(20\ \text{m/s})} = 2.4 \times 10^{-34}\ \text{m}.$$ ✔

This is 24 orders of magnitude less than the diameter of an atom!

(*b*) $$\lambda_{\text{de Broglie, electron}} = \frac{6.626 \times 10^{-34}\ \text{J}\cdot\text{s}}{(9.1 \times 10^{-31}\ \text{kg})(5.0 \times 10^{6}\ \text{m/s})} = 1.5 \times 10^{-10}\ \text{m} = 0.15\ \text{nm}.$$ ✔

Electron diffraction takes place on distances comparable to the spacing between atoms, whereas a baseball would diffract only through apertures more than 10^{24} times smaller than the typical spacing between atoms. (But the diameter of a baseball is about 0.10 m, so such an experiment is not possible.)

CONCEPTS

34.12 How would the electron diffraction pattern in Figure 34.23 change if the electrons were traveling more slowly?

34.5 Photons

We have now found that particles have wave properties, but these properties are observed only when the particle wavelength is comparable to the size of the objects the particles interact with. Up until now, you have most probably thought of light as being solely a wave. However, the dual wave-particle nature of particles may lead you to wonder whether light, too, is not just a wave but also a particle.

CONCEPTS

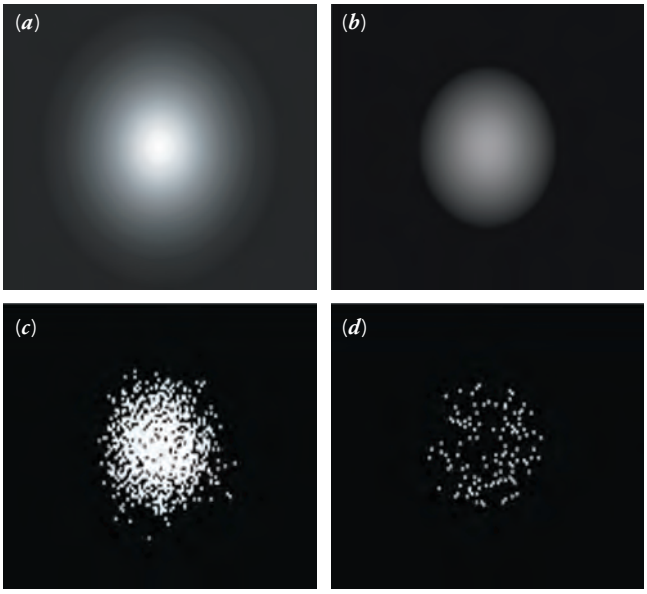
34.13 Compare the two-slit interference pattern obtained with electrons (Figure 34.25) with the two-slit interference pattern obtained with light (Figure 34.8). If light and electrons exhibit a similar wave-particle duality, how might you modify the two-slit experiment with light shown in Figure 34.6 to observe light behaving like a particle?

Figure 34.26*a* shows the image obtained by shining a beam of light onto the sensor chip of a digital camera. Pixels in the center of the beam register more light, and therefore the image of these pixels is brighter. The pixels at the edge of the beam are dimmer than those at the center, and outside the beam the pixels are black. If we place in front of the sensor chip a filter that lets only 50% of the light through, the image darkens because the brightness measured by each camera pixel is cut in half (Figure 34.26*b*).

Suppose we keep adding such filters, cutting the beam intensity in half with each addition. Does the image keep getting proportionally darker? If you carry out the experiment, you will discover that it does not, for one of two reasons. The first reason is mechanical: If you use the sensor chip of a digital camera, it stops detecting below a certain level. As you decrease the

*"de Broglie" is pronounced "duh-Br-uh-y," the "y" sounding as in "yikes."

Figure 34.26 Images formed by using the sensor of a digital camera to record increasingly faint beams of light for the same exposure period. The separate dots recorded for the faintest beams reveal the particle-like behavior of light.



intensity of the beam by adding more and more filters, the image first gets grainy and then turns black.

The second reason has to do with the fundamental nature of light. Even if you use an extremely sensitive detector, you will still see that once the beam becomes very weak, adding another filter does not simply halve the image intensity. Instead, as shown in Figure 34.26*c* and *d*, the image of the beam breaks up into individual point-like flashes of equal intensity, resembling the impacts of individual particles. The impacts may at first appear to be randomly distributed within the profile of the beam, but if you accumulate many of these impacts, you discover that the probability of observing a flash in a given location follows the intensity profile of the beam—the impacts are more likely to occur near the center of the beam. In fact, the probability of observing an impact in a particular location is proportional to the intensity of the beam at that location.

As the beam intensity is reduced, the individual impacts become separated in time. No two impacts occur at the same instant. Nor are any "half impacts" ever recorded. From these observations, we conclude that light indeed has particle properties as well as wave properties. The "particles" of light are called **photons.** As we discuss further in Section 34.10, a photon represents the basic unit of a light wave and carries a certain amount of light energy. For a photon of frequency f, this energy equals hf, where h is again Planck's constant. Photons thus represent the quantum of electromagnetic energy—they cannot be subdivided.

EXAMPLE 34.5 Photons from a light bulb

A 50-W incandescent light bulb emits about 5.0 W of visible light. (The rest is converted to thermal energy.) If a circular aperture 5.0 mm in diameter is placed 1.0 km away from the light bulb, approximately how many photons reach the aperture each second?

❶ **GETTING STARTED** This problem asks me to relate the power of light emitted by a light bulb to the rate at which photons pass through a certain area at a particular distance from the bulb. I can approximate the light bulb as a point source of light, meaning that its light is radiated uniformly in all directions. I can also simplify the problem by assuming that all of the light has the same wavelength, and I choose 500 nm for that wavelength. (This is a significant simplification because real light bulbs emit all visible wavelengths of light as well as infrared.)

❷ **DEVISE PLAN** I begin by determining the intensity—the power per unit area—of light produced by the light bulb over a sphere of radius 1 km, which tells me the intensity at the aperture. I can then multiply this intensity by the area of the aperture to obtain the power passing through the aperture, and I multiply that by 1 s to calculate the energy passing through the aperture in 1 s. Finally I shall use the relationship between photon energy and wavelength to determine the number of photons corresponding to that amount of energy.

❸ **EXECUTE PLAN** The intensity at the aperture is given by the power emitted by the bulb divided by the surface area of a sphere of radius 1.0 km:

$$\frac{(5.0\ \text{W})}{(4\pi)(1.0\times10^{3}\ \text{m})^{2}} = 4.0\times10^{-7}\ \text{J/m}^2\cdot\text{s},$$

and therefore the amount of energy passing through a circular hole of radius 2.5 mm in 1.0 s is

$$(4.0\times10^{-7}\ \text{J/m}^2\cdot\text{s})[\pi(2.5\times10^{-3}\ \text{m})^2] = 7.8\times10^{-12}\ \text{J}.$$

The energy of a single photon of wavelength 500 nm is

$$\begin{aligned} E = hf &= \frac{hc_0}{\lambda} \\ &= \frac{(6.626\times10^{-34}\ \text{J}\cdot\text{s})(3.00\times10^{8}\ \text{m/s})}{(500\times10^{-9}\ \text{m})} \\ &= 3.98\times10^{-19}\ \text{J}, \end{aligned}$$

and so the number of photons that corresponds to the amount of energy passing through the hole is

$$\frac{7.8\times10^{-12}\ \text{J}}{3.98\times10^{-19}\ \text{J}} = 2.0\times10^{7}\ \text{photons.}\ \checkmark$$

❹ **EVALUATE RESULT** The number of photons I obtain is great even though I know from experience that the amount of light entering my eye 1.0 km from a 50-W bulb is exceedingly small. However, photons contain a vanishingly

(Continued)

CONCEPTS

small amount of energy (about 10^{-19} J per photon), and so my answer is not implausible.
I assumed that all the photons have the same 500-nm wavelength. In reality the light bulb emits both longer- and shorter-wavelength photons. Suppose various wavelengths are equally distributed on both sides of 500 nm in the spectrum of the light emitted by the bulb. Photons with longer wavelengths have less energy and therefore there are more of them, but photons with shorter wavelengths have greater energy, and so there are fewer of them. The overall result is therefore about the same number of photons as the number I calculated using 500 nm as the only wavelength. To a first approximation, therefore, my result should be correct.

What are photons? How can light be a wave *and* consist of photons that behave like particles? I cannot answer this question because no one really knows the answer. I can, however, describe how photons behave. As you will see, their behavior defies common sense—it is unlike anything we ever experience.

If we reduce the intensity of a light beam (or a beam of any other type of electromagnetic radiation) so greatly that the flashes from the impacts of individual photons are well separated in time and then aim this beam at two very small, very closely spaced detectors (**Figure 34.27*a***), we see that each photon is detected by either one detector or the other. A simultaneous impact on both detectors is never recorded. (The detectors can be made as small as 1 μm wide and spaced by just a fraction of a micrometer.) This observation suggests that each photon takes a definite path—toward either detector A or detector B.

Now imagine replacing the two detectors in Figure 34.27*a* by two narrow slits of the same size as the detectors and placing a screen some distance back from the two slits (Figure 34.27*b*). The pattern on the screen initially looks like random impacts from photons that make it through either one slit or the other. As we accumulate the impacts of many photons, however, an interference pattern emerges (Figure 34.27*c*) that is identical to the one obtained by shining an intense beam of light on the two slits.

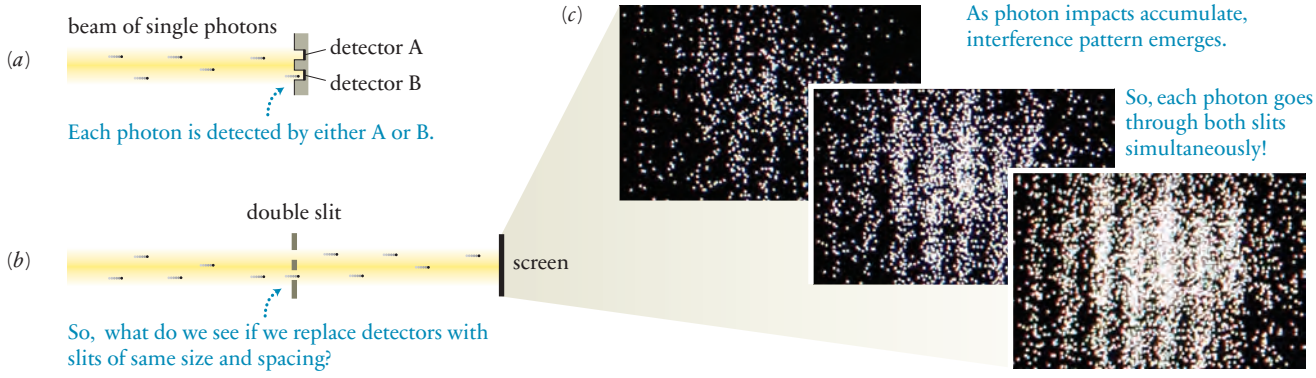
Notice that replacing the two detectors with two slits doesn't change anything about the beam or the photons contained in it. However, the experiment using the two detectors suggests that photons are particles detected by one or the other of the detectors; the experiment using the two slits indicates that each photon is a wave traveling through both slits simultaneously! So which is it? Do the photons travel through one slit only or through both? If we physically cover one of the slits (and therefore *force* the photons to go through the other slit), the interference pattern disappears. The pattern that emerges after accumulating many photons behind the slit corresponds to the diffraction pattern of light behind a single slit.

What this means is that photons behave as discrete particles when they are being detected. In transit, however, the wave nature of photons dictates their behavior. In other words, light behaves *both* like a wave and like a particle. It is impossible to explain the results of the above set of experiments by treating light as only a wave or only a particle; it must have qualities of both.

34.14 Figure 34.8 shows the interference pattern obtained by shining a strong laser beam on a pair of slits. If instead a very weak beam is shone on the same slits, so that the photons pass through the slits one photon at a time, at what angles is the probability of observing a particular photon the greatest?

CONCEPTS

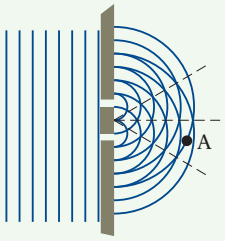
Figure 34.27 When we record a low-intensity beam of light with individual detectors, the beam acts like a stream of particles. Passing it through a double slit, however, causes an interference pattern to emerge.



Self-quiz

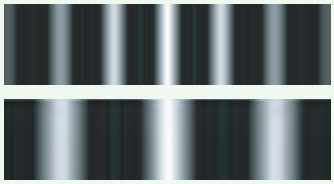
1. At point A in **Figure 34.28**, do the waves from the two slits add or cancel?

Figure 34.28



2. If the two sets of fringes shown in **Figure 34.29** were produced by the same diffraction grating, which set is the product of the longer-wavelength radiation?

Figure 34.29



3. Coherent light of wavelength λ is normally incident on two slits separated by a distance d. What is the greatest possible fringe order?
4. Consider a proton and an electron moving at the same speed. Which has the longer wavelength?
5. Given the relationship between the energy E of a photon and its frequency f and the de Broglie expression relating momentum $p = mv$ and wavelength λ, determine the ratio E/p for a photon.

Answers

1. At A, the crest of one wave overlaps the trough of the other wave, which means the waves cancel.
2. Because the wavelength is proportional to the sine of the angle the rays make with the normal to the diffraction grating, the fringes with the greater spacing were produced by the longer-wavelength radiation.
3. The fringe order is given by $d \sin\theta = \pm m\lambda$. Because the maximum value of $\sin\theta$ is 1, $d = m_{\max}\lambda$ and so the maximum fringe order is $m_{\max} = d/\lambda$. (Because $m_{\max}$ is an integer, you must truncate the value you obtain by dividing d by λ. For example, if $d/\lambda = 2.8$, then the greatest fringe order is 2.)
4. Because a proton has greater mass than an electron and because the de Broglie wavelength is inversely proportional to mass, $\lambda = h/mv$, the electron has a longer wavelength than the proton.
5. The relationship between the energy of a photon and its frequency is $E = hf$ (see Section 34.5). The de Broglie wavelength is given by $\lambda = h/mv = h/p$, so $p = h/\lambda$. Therefore $E/p = hf/(h/\lambda) = f\lambda = c$, where c is the speed of light.

34.6 Multiple-slit interference

Let us now calculate the interference pattern produced by an electromagnetic wave normally incident on a barrier pierced by multiple closely spaced, very narrow slits. The width of each slit is much less than the wavelength of the radiation, so each slit serves as a point source of radiation. We begin by determining the pattern created when the barrier has just two very narrow slits.

As discussed earlier, the point sources corresponding to the two slits are in phase. Therefore the electric fields of the two waves that reach the screen differ in phase only due to any difference in the distance each wave travels from its slit to the screen. We shall refer to this difference as the *path-length difference* δs (**Figure 34.30*a***). Taking the screen to be at position $x = 0$, we can write the electric fields of the waves traveling in the directions shown in Figure 34.30*a* as

$$E_1 = E_0 \sin \omega t \tag{34.1}$$

and

$$E_2 = E_0 \sin (\omega t + \phi), \tag{34.2}$$

where E_0 is the amplitude of the electric field and ϕ is a phase constant that is equal to the phase difference that results from the path-length difference between the two waves. The phase difference divided by 2π is the fraction of a cycle by which the two waves differ. This equals the path-length difference δs divided by the wavelength, or $\phi/2\pi = \delta s/\lambda$, and so

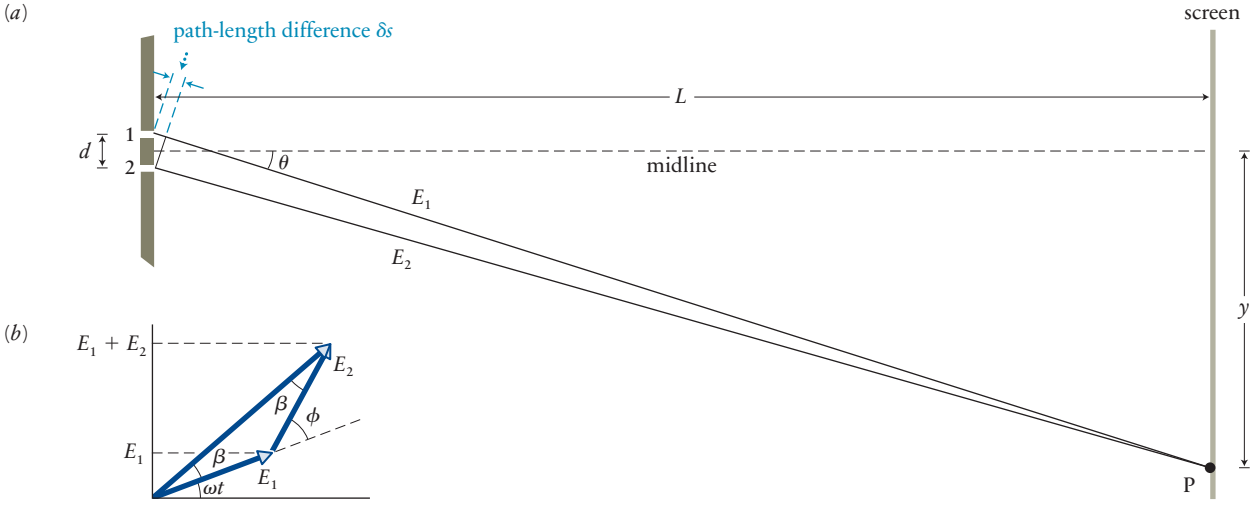
$$\phi = \frac{2\pi}{\lambda} d \sin \theta, \tag{34.3}$$

where d is the distance between the slits and θ is the angular position where the two rays meet on the screen.

We observe a bright fringe—a maximum in the intensity—when the two rays interfere constructively. This is the case when the phase difference ϕ equals an integer number times 2π:

$$\phi_m = \pm m(2\pi), \quad \text{for } m = 0, 1, 2, 3, \ldots. \tag{34.4}$$

Figure 34.30 (*a*) Interference of light diffracted by two very narrow slits. (*b*) We sum the phasors associated with the electric fields of the coherent light sources at slits 1 and 2.



Combining Eqs. 34.3 and 34.4 and solving for $\sin\theta$ determine the angles θ_m for which bright fringes occur:

$$\sin\theta_m = \pm\frac{m\lambda}{d}, \quad \text{for } m = 0, 1, 2, 3, \ldots. \quad \text{(bright interference fringes)}, \tag{34.5}$$

as we found in Section 34.2.

A dark fringe—a minimum in the intensity—occurs when the two rays interfere destructively. This is the case when the phase difference equals a odd number times π:

$$\phi_n = \pm(2n-1)\pi, \quad \text{for } n = 1, 2, 3, \ldots. \tag{34.6}$$

Substituting ϕ into Eq. 34.3 and solving for $\sin\theta$ determine the angles θ_n for which dark fringes occur:

$$\sin\theta_n = \pm\frac{(n-\frac{1}{2})}{d}\lambda, \quad \text{for } n = 1, 2, 3, \ldots \quad \text{(dark interference fringes)}. \tag{34.7}$$

To calculate the light intensity as a function of θ, we start by determining the sum of the electric fields E_1 and E_2 (Eqs. 34.1 and 34.2) using phasors. The amplitude of the sum of two sinusoidal quantities is equal to the vector sum of the phasors representing those two quantities (see Section 32.6). The phasors that represent E_1 and E_2 and their sum phasor are shown in Figure 34.30*b*. For the isosceles triangle formed by the E_1 and E_2 phasors, the exterior angle ϕ is equal to the sum of the two opposite interior angles β, and so $\beta = \frac{1}{2}\phi$. From this, we calculate the amplitude E_{12} of the sum of the two electric fields:

$$E_{12} = 2(E_0\cos\beta) = 2(E_0\cos\tfrac{1}{2}\phi). \tag{34.8}$$

Because this combined electric field oscillates at the same frequency as the incident wave, the time-dependent electric field at the screen is $E = E_{12}\sin\omega t$. At the central bright fringe, $\theta = 0$ and the two beams are in phase, so $\phi = 0$. The amplitude of E_{12} at the central bright fringe is therefore twice the amplitude of the incident wave, as we would expect.

We found in Chapter 30 (Eq. 30.36) that the intensity S of an electromagnetic wave is proportional to the product of the magnitudes of the electric and magnetic fields E and B: $S = EB/\mu_0$. Because B is proportional to E (Eq. 30.24, $E = Bc_0$), the intensity of light is commonly written in terms of E^2:

$$S = \frac{1}{\mu_0}EB = \frac{E(E/c_0)}{\mu_0}. \tag{34.9}$$

Substituting Eqs. 34.8 and 34.3 into Eq. 34.9 yields

$$S = \frac{4E_0^2}{\mu_0 c_0}\cos^2\left(\frac{\pi d\sin\theta}{\lambda}\right)\sin^2\omega t. \tag{34.10}$$

Visible electromagnetic waves oscillate at such high frequencies (10^{14} Hz to 10^{15} Hz) that we ordinarily measure the time-averaged intensity. The time average of $\sin^2\omega t$ is $\frac{1}{2}$. We can then write the time-averaged intensity of the interference pattern in terms of the time-averaged intensity of the

incident wave: $S_{0,\text{av}} = E_0^2/(2\mu_0 c_0)$. Thus substituting $\frac{1}{2}$ for $\sin^2 \omega t$ and $S_{0,\text{av}}(2\mu_0 c_0)$ for E_0^2 in Eq. 34.10 gives us

$$S_{\text{av}} = 4S_{0,\text{av}} \cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right). \tag{34.11}$$

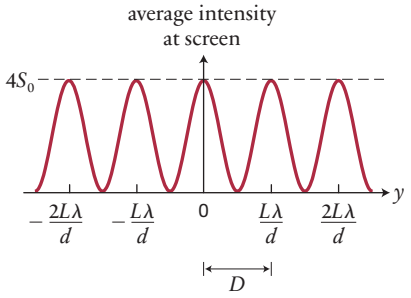
The maximum intensity of the interference pattern is *not* just the sum of the intensities of the two interfering waves—it is twice the sum! Because intensity is proportional to the square of the electric field, we must add the electric fields and then calculate the intensity from the square of the combined electric field.

34.15 How can the energy of the closed system made up of the two interfering waves remain constant (as the energy law states it must) if the maximum time-averaged intensity is four times the individual time-averaged intensities of the two waves?

Now let us work out how this pattern looks on the screen. If we consider only small angles θ, we can approximate $\sin\theta \approx \tan\theta = y/L$, where y is the position on the screen corresponding to the angle θ measured from the midline between the slits and L is the distance between the screen and the barrier that contains the slits, as shown in Figure 34.30*a*. (Positive y corresponds to positions above the midline; negative y to positions below the midline.) We can then write the time-averaged intensity as

$$S_{\text{av}} = 4S_{0,\text{av}} \cos^2(\phi/2) \approx 4S_{0,\text{av}} \cos^2\left(\pi d \frac{y}{L\lambda}\right). \tag{34.12}$$

Figure 34.31 Average intensity produced on a screen by two-slit interference.



The intensity varies periodically with the phase difference ϕ. Near the central maximum, where θ is small, the intensity also varies periodically with y (**Figure 34.31**).

What is the distance D between adjacent intensity maxima of this pattern? Substituting y/L for $\sin\theta$ in Eq. 34.5, we obtain, for the positions of the maxima corresponding to any two values m and $m+1$ of our order integer m,

$$m\lambda = d\frac{y_m}{L} \quad \text{and} \quad (m+1)\lambda = d\frac{y_{m+1}}{L}. \tag{34.13}$$

Subtracting the first of these equation from the second yields

$$\lambda = \frac{d}{L}(y_{m+1} - y_m). \tag{34.14}$$

The distance between adjacent maxima is $y_{m+1} - y_m = D$, so

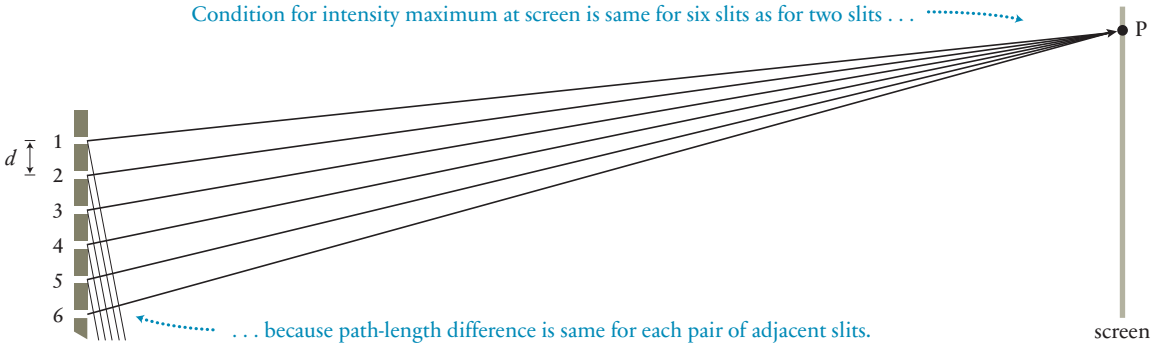
$$D = \frac{L}{d}\lambda. \tag{34.15}$$

This expression tells us that the distance between the maxima is proportional to the wavelength of the light. The interference pattern "magnifies" the wavelength by the factor L/d.

If there are more than two slits in the barrier, the analysis proceeds in the same fashion as before. **Figure 34.32** shows a wave diffracting through six slits, each separated from its immediate neighbors by distance d. The

QUANTITATIVE TOOLS

Figure 34.32 Interference of light diffracted by six narrow slits.



condition for a maximum in the intensity pattern is the same regardless of the number of slits because if the difference in path length for waves from slits 1 and 2 is λ, then the difference in path length for any pair of adjacent slits is also λ. The principal maxima therefore appear at those locations where

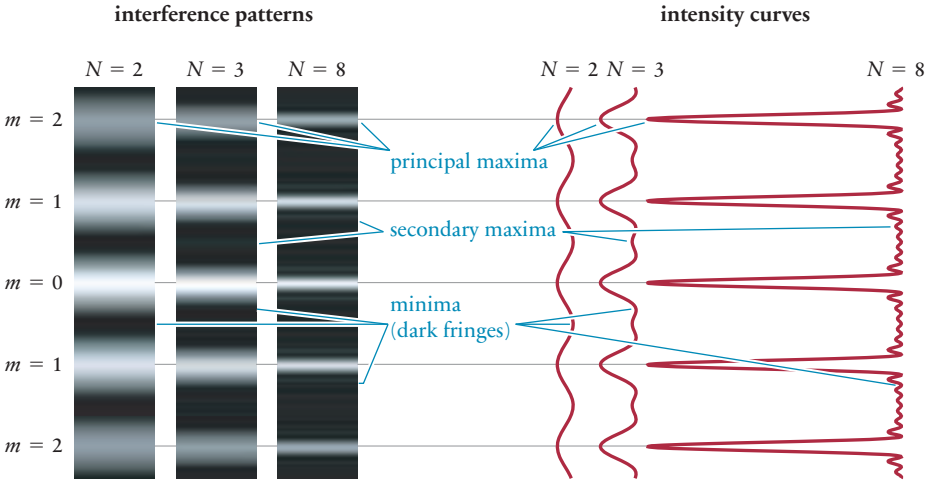
$$d \sin \theta_m = \pm m\lambda, \quad \text{for } m = 0, 1, 2, 3, \ldots \text{ (principal maxima)}, \tag{34.16}$$

just as we found for a two-slit barrier (Eq. 34.5).

The principal maxima occur at the same angles regardless of the number of slits. However, as discussed in Checkpoint 34.5, the *intensity* at the maxima increases with the number of slits. As we found in Section 34.2, as the number of slits increases, the interference fringes also become narrower (**Figure 34.33**). This is so because the minima closest to the m^{th} principal maximum occur for the ratio k/N that is as close as possible to m—namely, $(mN \pm 1)/N$. As the number of slits N becomes very great, the minima lie very close to m and so the interference pattern has extremely sharp maxima separated by broad dark regions with very faint secondary maxima.

As discussed in Section 34.2, an important use of diffraction gratings is to disperse (separate) light into its constituent wavelengths in order to better understand the light source. The amount of information that can be obtained from a spectrum depends on whether the wavelengths of interest can be distinguished from one another in the spectrum (such wavelengths are said to be *resolved*).

Figure 34.33 Interference pattern produced by gratings with two, three, and eight slits.



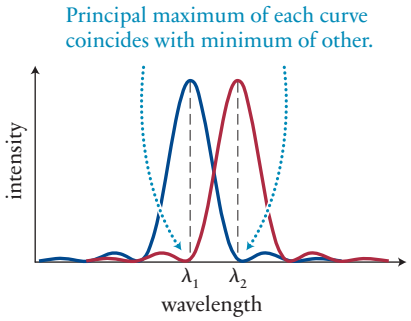


Figure 34.34 Two clearly separated bright fringes for light of different wavelengths.

Two wavelengths can be distinguished from each other if the principal maximum of one falls in the first dark region of the other, as shown in **Figure 34.34**. We found in Section 34.2 that, in general, a minimum occurs for interference through N slits when

$$d \sin \theta_{\text{min}} = \pm \frac{k}{N} \lambda, \quad \text{for integer } k \text{ that is not an integer multiple of } N \text{ (dark interference fringes).} \tag{34.17}$$

The angular position of the principal maxima increases with wavelength. Therefore, two wavelengths λ_1 and λ_2 can be distinguished from each other if the principal maximum for the longer wavelength falls at an angle greater than or equal to the angle for the $n = 1$ minimum for the shorter wavelength. As we first explored in Checkpoint 34.7, the separation between the slits is critical in determining the smallest wavelength difference that can be distinguished (this wavelength difference is called the *resolution*).

EXAMPLE 34.6 Resolution of wavelengths in a diffraction grating

An astronomer wishes to determine the relative heights of the intensity peaks for the bright fringes produced by two wavelengths of radiation emitted by sodium atoms. The wavelengths are 589.0 nm and 589.6 nm, and she uses a diffraction grating with 500.0 slits/mm to disperse the light collected by her telescope. (*a*) In which order are the intensity maxima for these two wavelengths farthest apart from each other: $m = 0$, $m = 1$, or $m = 2$? (*b*) If the part of the diffraction grating covered by the light is 4.000 mm wide, are the second-order principal maxima produced by these two spectral lines distinguishable from each other?

❶ **GETTING STARTED** This problem is about the overlapping diffraction patterns produced by two very similar wavelengths of light as the light passes through a diffraction grating. To answer the questions, I need to calculate the angular positions of the principal intensity maxima for each wavelength in three orders of diffraction. I also need to determine the angular position of the minimum adjacent to the second-order principal maximum for each wavelength, in order to determine whether the second-order principal maxima of the two wavelengths are distinguishable.

❷ **DEVISE PLAN** For part *a*, I can use Eq. 34.16 to locate the $m = 0$, 1, and 2 principal maxima for both wavelengths. For part *b*, I can locate the minimum adjacent to the second-order principal maximum for $\lambda = 589.0$ nm by applying the discussion following Eq. 34.17 and then compare that location with the location of the second-order principal maximum for $\lambda = 589.6$ nm.

❸ **EXECUTE PLAN** (*a*) I start by solving Eq. 34.16 for θ:

$$\theta_m = \pm \sin^{-1}\left(\frac{m\lambda}{d}\right). \tag{1}$$

This result indicates that the principal maxima are farthest apart for $m = 2$.

To calculate the positions of the principal maxima, I substitute the appropriate values of m, λ, and d in Eq. 1. The separation d between the slits is $1/(500 \text{ slits/mm}) = 2.000 \times 10^{-3}$ mm $= 2000$ nm. The $m = 0$ principal maximum occurs at $\theta = 0$ for any wavelength. The first-order principal maxima are located at

$$\theta_{589.0} = \pm \sin^{-1}\left(\frac{589.0 \text{ nm}}{2000 \text{ nm}}\right) = \pm 17.13°$$

$$\theta_{589.6} = \pm \sin^{-1}\left(\frac{589.6 \text{ nm}}{2000 \text{ nm}}\right) = \pm 17.15°,$$

and the second-order principal maxima are located at

$$\theta_{589.0} = \pm \sin^{-1}\left(\frac{2 \times 589.0 \text{ nm}}{2000 \text{ nm}}\right) = \pm 36.09°$$

$$\theta_{589.6} = \pm \sin^{-1}\left(\frac{2 \times 589.6 \text{ nm}}{2000 \text{ nm}}\right) = \pm 36.13°.$$

For these two wavelengths, the second-order principal maxima are the ones farthest apart from each other. ✔

(*b*) Using the information given in the discussion following Eq. 34.17, I can write that the condition for the minimum adjacent to the second-order principal maximum for $\lambda = 589.0$ nm is

$$d \sin \theta_{\text{min},589.0} = \pm \lambda \frac{mN + 1}{N},$$

where I have replaced k in Eq. 34.17 by $mN + 1$ as explained in the discussion preceding that equation. Solving this expression for $\theta_{\text{min},589.0}$ gives

$$\theta_{\text{min},589.0} = \pm \sin^{-1}\left(\frac{\lambda}{d} \frac{mN + 1}{N}\right).$$

A region of the grating 4.000 mm wide contains 2000 slits. Substituting in the preceding expression, I get that, for $\lambda = 589.0$ nm, the minimum adjacent to the second-order principal maximum is at

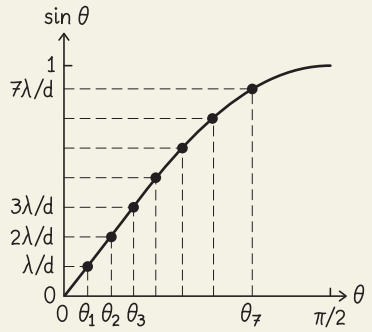
$$\theta_{\text{min},589.0} = \pm \sin^{-1}\left(\frac{589.0 \text{ nm}}{2000 \text{ nm}} \times \frac{4001}{2000}\right) = \pm 36.10°. \checkmark$$

QUANTITATIVE TOOLS

This minimum lies between 36.09°, the second-order principal maximum for $\lambda = 589.0$ nm, and 36.13°, the second-order principal maximum for $\lambda = 589.6$ nm, telling me that the second-order principal maxima for these two wavelengths are distinguishable from each other.

4 EVALUATE RESULT Equation 34.16 tells me that the angles at which principal maxima occur are equal to the inverse sine of integer multiples of λ/d. For small angles $\sin\theta_m \approx \theta_m$, and so the angles θ_m are approximately equally spaced (**Figure 34.35**). As the curve for $\sin\theta$ bends toward the horizontal, however, the distance between adjacent θ_m values increases, in agreement with the result I obtained.

Figure 34.35

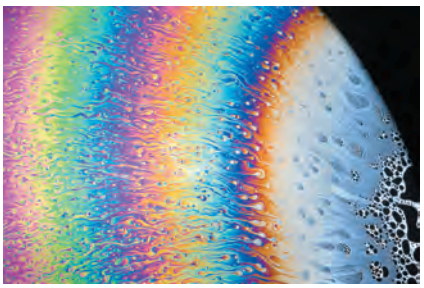


34.16 As the above discussion indicates, the separation distance between the principal maxima for different wavelengths increases as the fringe order m increases. Can you obtain an arbitrarily great separation distance by going to extremely high orders?

34.7 Thin-film interference

A familiar manifestation of the interference of light is the rainbow of colors reflected from thin films such as soap bubbles (**Figure 34.36**). This type of interference occurs in transparent materials whose thickness is comparable to the wavelengths of visible light. When such a thin material (we refer to it as a *film*) is either suspended in air, as in a soap bubble, or supported on a much thicker material with a different index of refraction, as in an oil slick on a puddle of water, white light reflecting from both the front and back surfaces of the film interferes, in the same manner as x rays reflecting from adjacent layers of atoms in a crystal. This is shown schematically in **Figure 34.37**. The film thickness t and index of refraction n_b, together with the angle of incidence, determine the path-length difference and corresponding phase difference between the reflected beams, and therefore determine which colors undergo constructive interference and which undergo destructive interference in any given direction.

Figure 34.36 Soap bubble.



To identify the conditions for constructive and destructive interference, we begin by expressing the electric fields of the waves reflected from the two surfaces, just as we did for waves passing through two slits in the preceding section:

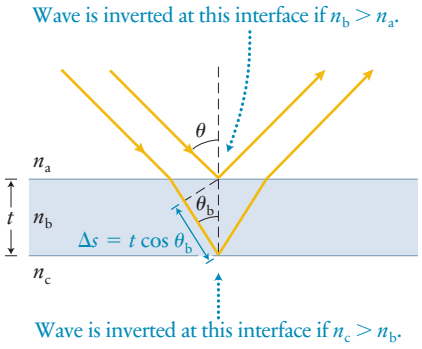
$$E_1 = E_0 \sin(\omega t) \tag{34.18}$$

$$E_2 = E_0 \sin(\omega t + \phi). \tag{34.19}$$

The phase difference ϕ is due to the path-length difference $2\Delta s$ and the effect of the reflections on the phases of each wave. The path-length difference gives rise to a phase difference

$$\phi_{\text{path}} = \frac{2\pi\Delta s}{\lambda_b} = \frac{2\pi(2t\cos\theta_b)}{\lambda/n_b} = \frac{4\pi n_b t\cos\theta_b}{\lambda}, \tag{34.20}$$

Figure 34.37 Thin-film interference.



QUANTITATIVE TOOLS

where the path-length difference Δs and the angle θ_b of the ray relative to the surface normal inside the film are shown in Figure 34.37, and $\lambda_b = \lambda/n_b$ is the wavelength of light inside the film (see Eq. 33.3). Because of refraction, expressing Δs in terms of the angle of incidence θ, instead of the angle in the film θ_b, involves Snel's law and produces a rather complicated result. However, for normal incidence $\cos \theta_b = 1$, and so the expression is greatly simplified.

To determine the phase difference due to the reflections from the two film surfaces, recall the discussion of mechanical waves at boundaries from Section 16.4. We saw there that when a wave pulse is launched in a first medium and travels into a second medium, the pulse is partially reflected and partially transmitted at the boundary. The reflected pulse is inverted relative to the incident pulse if the wave speed c_1 in the first medium is greater than the speed c_2 in the second medium. When $c_1 < c_2$, the incident pulse is not inverted upon reflection.

Let us now extend these ideas to sinusoidal electromagnetic waves. When the incident wave in medium 1 reflects at the boundary with medium 2 of greater index of refraction ($n_1 < n_2$), the speed is less in medium 2 ($c_2 < c_1$), and so the wave is inverted upon reflection. This inverting of the wave is equivalent to the wave undergoing a phase shift of π upon reflection. If instead $n_1 > n_2$, the wave is not inverted and there is no phase shift upon reflection.

A phase shift of π can occur at each of the two interfaces, depending on the refractive indices n_a, n_b, and n_c of the three media, as indicated in Figure 34.37. Thus the phase difference associated just with the reflections for the two reflected waves is

$$\phi_r = \phi_{r2} - \phi_{r1}, \tag{34.21}$$

where ϕ_{r1} and ϕ_{r2} are either π or 0, depending on the values of the indices of refraction. Therefore the total phase difference between the two reflected waves is

$$\phi = \frac{4\pi n_b t \cos \theta_b}{\lambda} + \phi_{r2} - \phi_{r1}. \tag{34.22}$$

Because ϕ_{r1} and ϕ_{r2} must each be π or 0, the effect of the two reflections on the phase cancels out when $\phi_{r1} = \phi_{r2}$, and the phase difference is due entirely to ϕ_{path}. If ϕ_{r1} is not equal to ϕ_{r2}, then ϕ differs from ϕ_{path} by π, causing the constructive and destructive interference conditions to be switched from what they would be due to just ϕ_{path}.

In our discussion of x-ray diffraction, the light source consisted of monochromatic x rays, and we found that constructive interference between adjacent planes takes place at only certain angles. In the current context—light reflections from thin films—we usually deal with light across the visible spectrum. Therefore ϕ depends on both wavelength and angle of incidence. For normal incidence, Eq. 34.22 simplifies to

$$\phi = \frac{4\pi n_b t}{\lambda} + \phi_{r2} - \phi_{r1} \quad \text{(normal incidence)}. \tag{34.23}$$

Then as before, constructive interference occurs when the phase difference corresponds to an integer number times 2π, and destructive when it corresponds to an odd number times π.

QUANTITATIVE TOOLS

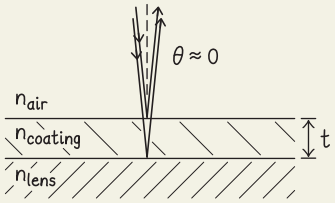
EXAMPLE 34.7 Antireflective coating

Eyeglass lenses made of crown glass ($n = 1.52$) are given a thin coating of magnesium fluoride ($n = 1.38$) to minimize reflection of light from the lens surface. What is the minimum coating thickness for which reflection from the lens surface is minimized?

❶ **GETTING STARTED** Reflection is minimized at all values of the coating thickness that cause the waves reflected from the front coating surface to interfere destructively with those reflected from the back coating surface. My task therefore is to obtain the minimum thickness needed for destructive interference.

❷ **DEVISE PLAN** I see from Eq. 34.22 that the condition for destructive interference depends not only on the coating thickness t and index of refraction n_b but also on two variables I have no values for: the angle of incidence and the wavelength of the light. Thus I must make some simplifying assumptions to solve this problem. As in the text discussion, I shall consider only light normally incident on the coated lens and shall assume the lens and coating surfaces are flat, as shown in my sketch (**Figure 34.38**). I choose a representative visible wavelength, 500 nm—in the middle of the visible spectral range.

Figure 34.38



To obtain the condition for destructive interference, I can use Eq. 34.23 to determine the phase difference ϕ at normal incidence and equate it to an odd number times π radians. To work out the phase shifts ϕ_{r1} and ϕ_{r2} that occur at each reflection, I note that at the air-coating interface the light reflects from a medium for which the index of refraction is greater than that of the medium through which the incident wave travels, and so this wave undergoes a π phase shift upon reflection: $\phi_{r1} = \pi$. The wave that reflects from the coating-lens interface likewise reflects from a medium for which the index of refraction is greater than that of the medium through which the incident wave travels and thus also undergoes a π phase shift ($\phi_{r2} = \pi$). Consequently, ϕ_{r1} and ϕ_{r2} in Eq. 34.23 cancel, leaving only the phase difference due to the path-length difference. The minimum thickness produces a phase difference corresponding to the smallest number of cycles that gives destructive interference—namely, half a cycle.

❸ **EXECUTE PLAN** To determine the thickness that produces destructive interference for 500 nm light, I express the phase difference using Eq. 34.23 and equate it to a half-cycle phase difference:

$$\frac{4\pi n_b t}{\lambda} + \phi_{r2} - \phi_{r1} = \pi.$$

Substituting $\phi_{r1} = \phi_{r2} = \pi$, solving for the thickness t, and substituting values give

$$t = \frac{\lambda}{4n_b} = \frac{500\text{ nm}}{4(1.38)} = 90.6\text{ nm}. \checkmark$$

❹ **EVALUATE RESULT** The coating is thinner than the wavelength of the light traveling through the coating. That makes sense because, to create destructive interference, the wave that travels through the coating and reflects from the coating-lens interface must travel only half a wavelength farther than the wave that reflects from the air-coating interface.



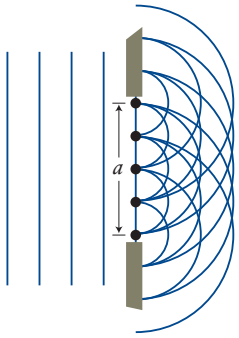
34.17 When oil spreads on water, bands of different colors are visible. What causes the different colors?

34.8 Diffraction at a single-slit barrier

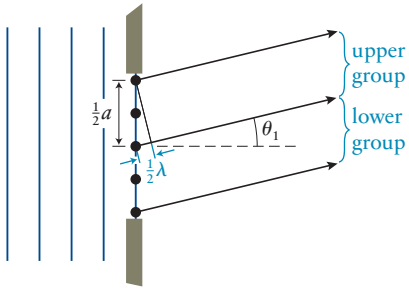
Let us now describe quantitatively the diffraction by a single slit. Huygens' principle (see Section 17.4) allows us to describe the wavefront that reaches the opening as a series of point sources that emit spherical wavelets (**Figure 34.39*a*** on the next page). The wave beyond the slit is the superposition of all these spherical wavelets. In the original direction of propagation, all of the waves add in phase, and as a result the amplitude of the transmitted wave is the maximum possible: the sum of all the individual wave amplitudes. If we divide the rays representing the wavelets into two equal groups, as shown in Figure 34.39*b*, we can pair each ray in the upper half with a corresponding ray in the lower half. In Figure 34.39*b*, for example, we can pair the top ray in the upper group with the top ray in the bottom group. As we discussed in Section 34.2, these rays are essentially parallel when they emerge from the slits; as they travel and intersect at some

Figure 34.39 Diffraction through a narrow slit.

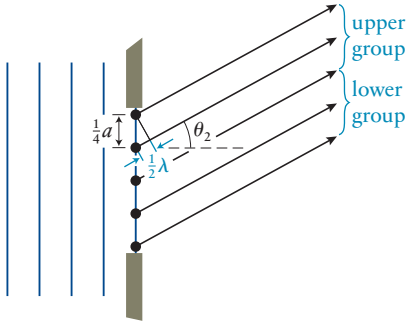
(*a*) Wavefront in slit treated as a series of point sources



(*b*) Condition for first-order dark fringe: path lengths differ by $(\frac{1}{2}a)\sin\theta_1 = \lambda/2$



(*c*) Condition for second-order dark fringe: path lengths differ by $(\frac{1}{4}a)\sin\theta_2 = \lambda/2$



location on a screen placed far to the right of the slits, the rays must travel different distances to reach that location. For rays traveling at an angle θ to the original propagation direction, the difference in path length from two such corresponding rays is $(a/2)\sin\theta$, where a is the width of the aperture. When the path lengths differ by half a wavelength, $(a/2)\sin\theta_1 = \lambda/2$, the rays interfere destructively, which means that the direction for the first-order dark fringe in a single-slit diffraction pattern is given by

$$\sin\theta_1 = \frac{\lambda}{a} \quad \text{(first-order dark diffraction fringe).} \tag{34.24}$$

Dividing the rays into four groups (Figure 34.39*c*) leads to a dark fringe (that is, a minimum in transmitted intensity) when $(a/4)\sin\theta = \frac{1}{2}\lambda$. Thus the direction for the second-order dark fringe is given by

$$\sin\theta_2 = 2\frac{\lambda}{a} \quad \text{(second-order dark diffraction fringe).} \tag{34.25}$$

The general condition for a dark fringe is thus

$$\sin\theta_n = \pm n\frac{\lambda}{a}, \quad n = 1, 2, 3, \ldots$$
$$\text{(dark diffraction fringes).} \tag{34.26}$$

Positive values of $\sin\theta_n$ correspond to dark fringes above the midline, while negative values correspond to dark fringes below the midline.

Calculation of the detailed intensity pattern is beyond the scope of this text, but we can examine some of the details of a representative pattern (**Figure 34.40**). In this single-slit diffraction pattern, the intensity of the first-order ($m = 1$) bright fringe is less than 5% of the intensity of the central ($m = 0$) bright fringe; the intensity of the second-order ($m = 2$) bright fringe is less than 2% of the central bright fringe intensity. In other words, most of the transmitted energy falls within the central peak. In general, in a single-slit diffraction pattern, the intensity is greatest at the central bright fringe and decreases rapidly with distance from the center of the pattern.

What if we want to calculate the linear positions of the dark diffraction fringes on a screen located a distance L away from a barrier containing a single slit (**Figure 34.41**) rather than the angular positions? For dark fringes located at small angles θ_n from the original direction of wave propagation, we can approximate $\tan\theta_n$ as $\sin\theta_n$:

$$y_n = L\tan\theta_n \approx L\sin\theta_n, \tag{34.27}$$

and so, from Eq. 34.26,

$$y_n = \pm n\frac{\lambda L}{a} \quad \text{(dark diffraction fringes).} \tag{34.28}$$

Figure 34.40 Intensity plot for a single-slit diffraction pattern.

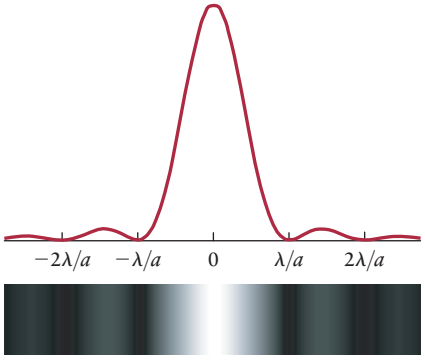
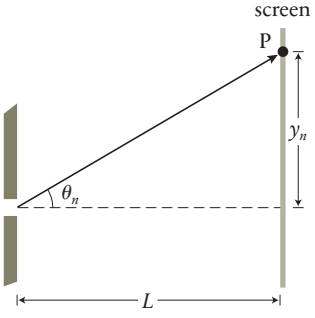


Figure 34.41 Calculating the positions of the dark fringes of a single-slit diffraction pattern.



QUANTITATIVE TOOLS

EXAMPLE 34.8 Spreading light

Consider the diffraction pattern shown actual size in Figure 34.40. If the pattern was formed by light from a 623-nm (red) laser passing through a single narrow slit and the screen on which the pattern was cast was 1.0 m away from the slit, what is the slit width?

❶ **GETTING STARTED** This problem asks me to relate the diffraction pattern in Figure 34.40, which was produced by a setup such as the one shown in Figure 34.41, to the width of the slit that produced it. Thus I need to relate the slit width to fringes whose position I can calculate from the parameters of the problem and can also measure on the image. Because the only variable I know how to calculate is the positions of minima in the diffraction pattern, I can measure the distance between the two first-order minima and relate that distance to the slit width and the geometry of the setup.

❷ **DEVISE PLAN** The positions of the two first-order minima, in terms of wavelength λ, slit-to-screen distance L, and slit width a, are given by Eq. 34.28 with $n = 1$. Subtracting the two values I obtain from each other gives me an expression for the distance between these two minima in terms of λ, L, and a. I am given the values of λ and L, and my task is to determine a. Thus if I know the distance between the n_1 minima, I can calculate a. Because the image in Figure 34.40 is actual size, I can measure this distance directly. Then I can solve Eq. 34.28 for a and insert my known values.

❸ **EXECUTE PLAN** Substituting $n = 1$ into Eq. 34.28 gives the linear positions of the two first-order minima:

$$y_1 = \pm\frac{\lambda L}{a},$$

so the distance w between the two minima is

$$w = \frac{2\lambda L}{a}. \tag{1}$$

I measure the distance between the centers of the two dark fringes on either side of the central bright fringe in Figure 34.40 to be 20 mm. Solving Eq. 1 for the slit width thus gives me

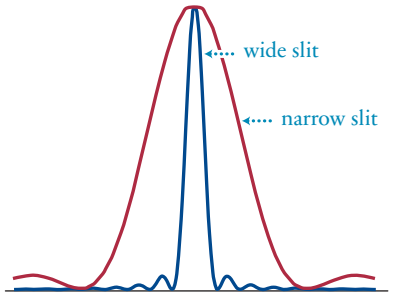
$$a = \frac{2\lambda L}{w} = \frac{2(623\times10^{-9}\text{ m})(1.0\text{ m})}{20\times10^{-3}\text{ m}}$$
$$= 6.2\times10^{-5}\text{ m} = 0.062\text{ mm} \quad ✔$$

❹ **EVALUATE RESULT** The slit width is about a factor of 100 greater than the wavelength of the light, and that ratio is consistent with the general range of slit sizes that produce noticeable diffraction of visible light.

As long as λ is small relative to the slit width a, so that the small-angle approximation for θ_n is valid, most of the diffracted light intensity falls within this region defined by Eq. 1 in Example 34.8. Note that w increases with decreasing a: The narrower the slit, the more the wave spreads out after passing through the slit (**Figure 34.42**). If the slit width is equal to or less than the wavelength of the light, there are no dark fringes, as we found in Checkpoint 34.16. The wave simply spreads out in all directions behind the slit, which means the slit behaves as a point source.

34.18 Using Eq. 34.24, calculate the angle at which the first dark fringe occurs when (*a*) $a < \lambda$ and (*b*) $a \gg \lambda$. Interpret your results.

Figure 34.42 The width of the central maximum in a diffraction pattern decreases as the slit is widened.



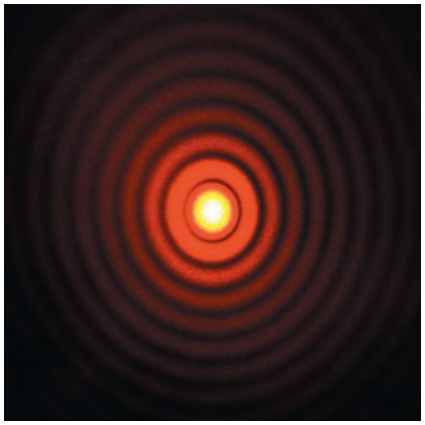
34.9 Circular apertures and limits of resolution

When light passes through a circular aperture, the symmetry of the aperture causes the resulting diffraction pattern also to be circular (**Figure 34.43**). A circular central bright fringe is surrounded by circular dark diffraction fringes and additional diffraction bright fringes. The central bright fringe is called the *Airy disk*, after the British astronomer and mathematician Sir George Airy, who developed the first detailed description of diffraction in 1835. Calculation of the circular diffraction pattern is rather involved, so all I shall do here is state the location of the first dark fringe. When light is diffracted by a circular aperture of diameter d, the first dark fringe occurs at angle θ_1 given by

$$\sin\theta_1 = 1.22\,\frac{\lambda}{d}. \tag{34.29}$$

The result is similar to the one obtained in Eq. 34.24 for a slit of width a, except that now the sine of the angle is increased by a factor of 1.22.

Figure 34.43 Diffraction pattern of light passing through a circular aperture.



The increase in the angle for the same wavelength and aperture size can qualitatively be understood as follows: Equation 34.24 is obtained by considering the interference between wavelets coming from a slit whose width is a everywhere. A circular aperture has width d only across a diameter; the rest of the aperture is narrower. As the aperture gets narrower, diffraction through it becomes more pronounced. The factor of 1.22 quantitatively accounts for the varying horizontal width of the circular aperture.

The angular size of the Airy disk given in Eq. 34.29 determines the minimum angular separation of two point sources that can be distinguished by observing them with a (circular) lens—*regardless* of the magnification of the lens! To understand what this means, imagine imaging two distant, closely spaced point sources with a lens. These sources could be anything, from stars to organelles in a biological cell. The sources are not coherent, and so we can consider the Airy disks formed by the light from each source separately without considering interference between sources.

If there is overlap in the Airy disks of the images observed through the lens, it is difficult to tell whether there are two point sources or just one. Two objects being observed through a lens are just barely distinguishable when the center of one diffraction pattern is located at the first minimum of the other diffraction pattern. This happens when the angular separation between the two objects is at least the angle given in Eq. 34.29. If this is the case, we say that the two objects are *resolved*. This condition for distinguishability is called **Rayleigh's criterion.**

Because the diameter d of the lens is always much greater than the wavelength of the light, the angle in Eq. 34.29 is always small. Thus the minimum angular separation θ_r for which two sources can be resolved is approximately equal to the sine of the angle

$$\theta_r \approx \sin \theta_r = 1.22 \frac{\lambda}{d}. \tag{34.30}$$

Two objects that are separated by an angle equal to or greater than θ_r satisfy Rayleigh's criterion. For this reason, the closest two objects can be to each other and still be distinguished with an optical instrument such as a microscope or telescope depends not on the magnification but on the wavelength of the light and the size of the smallest aperture in the instrument.

Figure 34.44 The resolution of these two stars improves as the aperture is made larger.

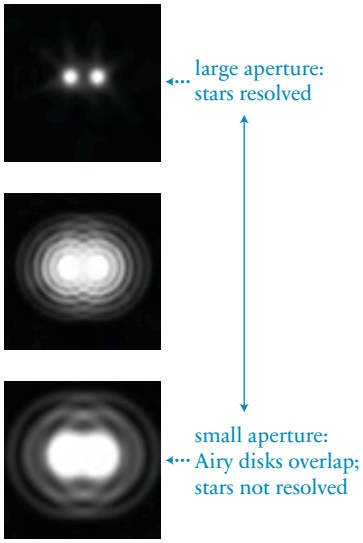


Figure 34.44 shows the images of two stars obtained with a telescope. An aperture placed in front of the lens shows the effects of diffraction. When the opening of the aperture is small (bottom image in Figure 34.44), the images of the two stars are merged—the two stars cannot be resolved. As the aperture is opened, θ_r decreases and the two images separate cleanly.

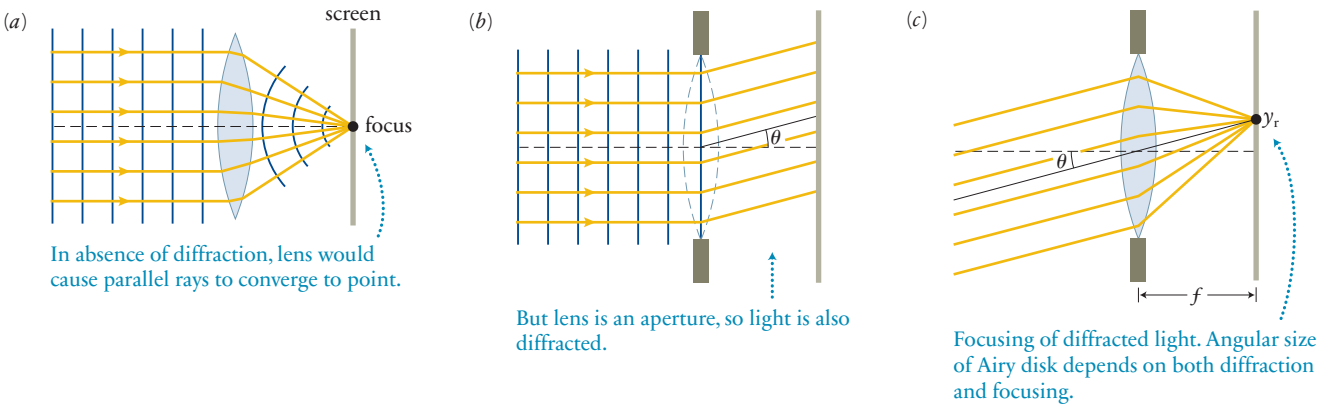
Diffraction also determines the linear size of the images of point sources. In Chapter 33, we stated that the image of a point source formed on a screen by a lens is a point. **Figure 34.45*a*** shows parallel rays from a distant point source focused by a lens onto a screen placed in the focal plane of the lens. Without diffraction, the image formed by these rays would be an infinitesimally small point. In fact, because the aperture through which the light passes—the lens—has a finite diameter, such rays do not focus to an infinitely small point, and the image formed is a diffraction pattern just like that shown in Figure 34.43. The angular size of the central bright fringe of this diffraction pattern—in other words, the Airy disk—is given by Eq. 34.29.

To calculate the radius of the Airy disk formed in the focal plane of the lens,* we must account for both diffraction and focusing. Figure 34.45*b*

*The radius of the Airy disk is smallest when the screen is in the focal plane of the lens, and it increases as the screen is moved closer to or farther from the lens.

QUANTITATIVE TOOLS

Figure 34.45 Analyzing the diffraction limit of a lens.



shows the diffraction of light through an aperture of the same diameter as the lens. Light that originally traveled parallel to the lens axis is diffracted by an angle θ. Now consider how the lens focuses the diffracted light. Parallel rays that make an angle θ with the lens axis are focused at a point located a distance y above the axis (Figure 34.45*c*, see also Figure 33.27). This distance y is given by $y = f \tan\theta$, where f is the focal length of the lens.

Our next step in determining the Airy disk radius is to calculate the distance y_r from the center of the disk to the first dark fringe in the diffraction pattern, which is found at the angle given by Eq. 34.29. In the small-angle limit, $\sin\theta \approx \tan\theta$ and so we can substitute $\sin\theta_r = y_r/f$ into Eq. 34.29, giving

$$y_r = 1.22\,\frac{\lambda f}{d}. \tag{34.31}$$

This expression gives the radius y_r of the Airy disk and the minimum size of the area to which light can be focused with light of wavelength λ by a lens of focal length f and diameter d. The best ratio of f/d that can be achieved with a lens is approximately equal to one, and so the smallest diameter to which light can be focused is about 2.5λ. This means that the smallest diameter of "points" in the resulting image is also 2.5λ.

This diffraction-determined minimum size of the features in an image is commonly called the *diffraction limit*. An industry in which the diffraction limit poses a serious problem is the manufacture of integrated circuits, such as computer chips. The transistors and logic gates described in Section 32.4 are produced by a series of processes known as *photolithography*, in which the semiconductor substrate is coated with a polymer that is sensitive to ultraviolet light. The polymer is then exposed to ultraviolet light in the pattern of the desired metal electrodes. This pattern of light is produced by illuminating a metal mask with holes in the shape of the electrodes and then imaging the resulting pattern of light onto the surface. Finally, the exposed polymer is dissolved with a chemical rinse, leaving the semiconductor surface exposed where metal is desired. The metal is then deposited in a subsequent step.

The smallest electrode that can be made by this process is therefore determined by the diffraction limit for the ultraviolet light ($\lambda \leq 150$ nm) used to produce the pattern. With ordinary optical technology, this requires electrodes to be at least 300 nm wide. Many researchers are searching for other ways to produce electrodes that are not limited by diffraction.

EXERCISE 34.9 A point is not a point

A magnifying glass that has a focal length of 0.25 m and a diameter of 0.10 m is used to focus light of wavelength 623 nm. (*a*) What is the radius of the smallest Airy disk that can be produced by focusing light with this lens? (*b*) How large is the Airy disk formed when this lens focuses a laser beam that has a 2.0-mm diameter?

SOLUTION (*a*) The radius of the smallest Airy disk is given by Eq. 34.31 with d equal to the diameter of the lens:

$$y = 1.22\frac{(623\times10^{-9}\text{ m})(0.25\text{ m})}{0.10\text{ m}} = 1.9\times10^{-6}\text{ m} = 1.9\ \mu\text{m}.$$ ✔

The minimum spot size is about three times greater than the wavelength of the light.

(*b*) In this case the radius is given by Eq. 34.31 with d equal to the diameter of the laser beam. I must use the beam diameter because the beam does not make use of most of the area of the lens but effectively defines its own aperture. I therefore have

$$y = 1.22\frac{(623\times10^{-9}\text{ m})(0.25\text{ m})}{0.0020\text{ m}} = 9.5\times10^{-5}\text{ m} = 95\ \mu\text{m}.$$ ✔

This Airy disk is much bigger than the one found in part *a*. This Airy disk radius means the central bright fringe has a diameter of 190 μm = 0.19 mm, which is smaller than the original beam diameter due to the focusing of the lens. Because the laser beam is much smaller than the diameter of the lens, it cannot be effectively focused by this lens. The diffraction partially cancels the focusing.

EXAMPLE 34.10 Blurry images

The width of one pixel in the sensor for a digital camera is about 2.0 μm. If the camera lens has a diameter of 40 mm and a focal length of 30 mm, is the resolution of the resulting image limited by the lens or the sensor?

❶ **GETTING STARTED** This problem involves comparing, for the image formed by a digital camera, the resolution limit due to diffraction and the limit due to the size of the pixels in the sensor. The limit due to diffraction is the size of the image of a point source; the limit due to the sensor is the width of a single pixel. The greater limit determines the image resolution. The problem does not specify the wavelength of the light involved, but because the problem is concerned with forming images with visible light, I choose λ = 500 nm, near the center of the visible spectrum.

❷ **DEVISE PLAN** The diffraction-limited image of a point source is the Airy disk at the center of the diffraction pattern, so I can use Eq. 34.31 to obtain the radius of the Airy disk formed by the camera. I can then compare the diameter (not the radius) of that disk with the width of a pixel. Whichever is greater limits the resolution possible for the image.

❸ **EXECUTE PLAN** Substituting the values given into Eq. 34.31, I obtain for the Airy disk radius

$$y = 1.22\frac{\lambda f}{d} = 1.22\frac{(0.500\times10^{-6}\text{ m})(30\times10^{-3}\text{ m})}{40\times10^{-3}\text{ m}} = 0.46\times10^{-6}\text{ m} = 0.46\ \mu\text{m}.$$

The diameter of the Airy disk is thus $2y$ = 0.92 μm. This is significantly less than the pixel width, which means the resolution of the image is limited by pixel width rather than by diffraction. ✔

❹ **EVALUATE RESULT** My result for the radius of the Airy disk is reasonable because typically the diffraction-limited width of the image of a point source is comparable to the wavelength of light emitted by the source. (Note that if I had chosen λ = 700 nm, the longest visible wavelength, the Airy disk radius would increase only by a factor of 700/500 = 1.4 and thus its diameter, 1.3 μm, would still not exceed the pixel width. If I had chosen a wavelength shorter than 500 nm, the Airy disk diameter would be less than my calculated value. So my conclusion is the same for any visible wavelength: The resolution is limited by the pixel width.)



34.19 Which of these three lenses offers (*a*) the highest resolution and (*b*) the lowest resolution: (*i*) f = 10 mm, d = 8 mm; (*ii*) f = 15 mm, d = 10 mm, (*iii*) f = 20 mm, d = 18 mm?

34.10 Photon energy and momentum

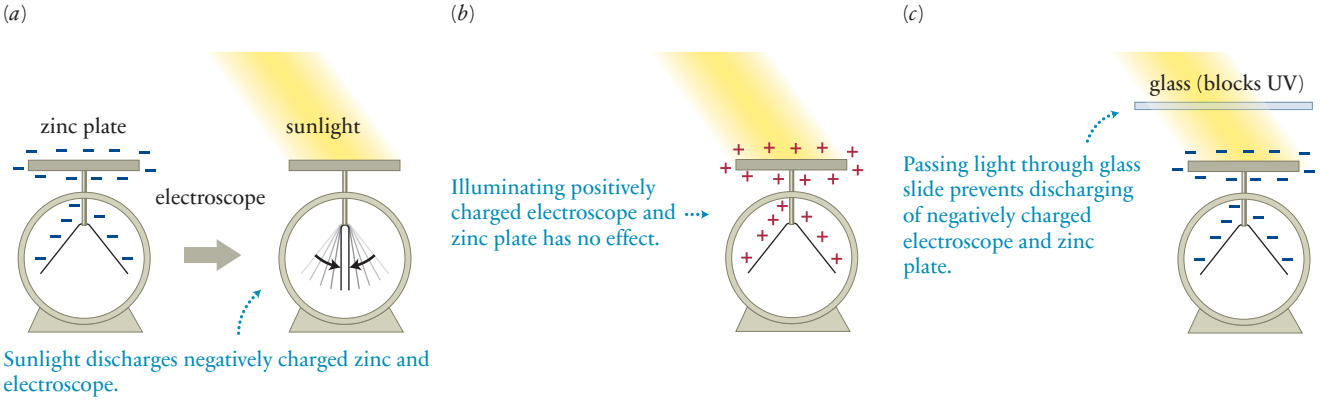
In Chapter 30 we described light as an electromagnetic wave that has a wavelength λ and a frequency f and moves in vacuum at speed c_0, such that

$$c_0 = \lambda f. \tag{34.32}$$

We also found that the energy density in the electromagnetic wave is proportional to the square of the amplitude of the electric field oscillation. In

QUANTITATIVE TOOLS

Figure 34.46 The photoelectric effect.



addition, the experiment described in Section 34.4 suggests that light has particle properties. If light always propagates at speed c_0 in vacuum, what determines its energy? And if light is a particle that has energy, shouldn't it also have momentum?

The answer to the first question is provided by the **photoelectric effect,** a surprising phenomenon that cannot be explained by thinking of light as a wave (**Figure 34.46**). Place a piece of metal, such as zinc, on an electroscope (see Section 22.3) that is negatively charged, as shown in Figure 34.46*a*. Some of the charge immediately moves to the zinc so that it, too, is negatively charged. If you then shine sunlight on the metal, the light discharges the zinc and the electroscope. If the electroscope is positively charged, however, as in Figure 34.46*b*, nothing happens when light shines on it. If we place a piece of glass in the beam of light (Figure 34.46*c*), nothing happens even if the zinc plate is negatively charged and the light is very intense.

What is going on? In the situations of Figures 34.46*a* and *b*, the light knocks electrons out of the zinc plate. When the plate is initially negatively charged, like-charge repulsion causes the ejected electrons to accelerate away from the plate. When the plate is initially positively charged, opposite-charge attraction causes the ejected electrons to be attracted back to the plate, so that the charge on the plate does not change. Ultraviolet radiation cannot pass through ordinary glass, and thus we conclude from the situation in Figure 34.46*c* that ultraviolet light is essential in order for electrons to be ejected.

The apparatus illustrated in **Figure 34.47** is used to study the photoelectric effect. It allows us to measure the energy of the ejected electrons while separately controlling either the wavelength or the intensity of the light. A zinc target T is placed in an evacuated quartz bulb (quartz is transparent to ultraviolet light), along with another metal electrode called the collector (C). A power supply is used to maintain a constant potential difference V_{CT} between the target and the collector. The current from the target to the collector is measured with an ammeter. If the target is kept at a lower potential relative to the collector, so that V_{CT} is negative, any electrons ejected from the target by the light are accelerated by the electric field and move to the collector. With negative V_{CT}, the current measured is proportional to the intensity of the light source, suggesting that ejecting each electron requires a certain amount of light energy.

If the potential difference V_{CT} is made slightly positive (so that the target is at a higher potential relative to the collector), there is a small current detected, but the electric field between T and C now *slows down* any ejected electrons that initially move toward the collector. As V_{CT} increases, there is a certain value of V_{CT} at which the flow of electron stops completely, as

Figure 34.47 Apparatus to study the photoelectric effect. The potential difference V_{CT} is positive when $V_T > V_C$.

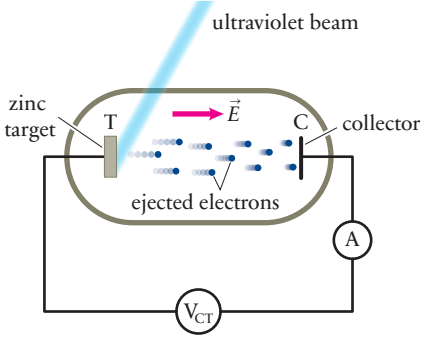
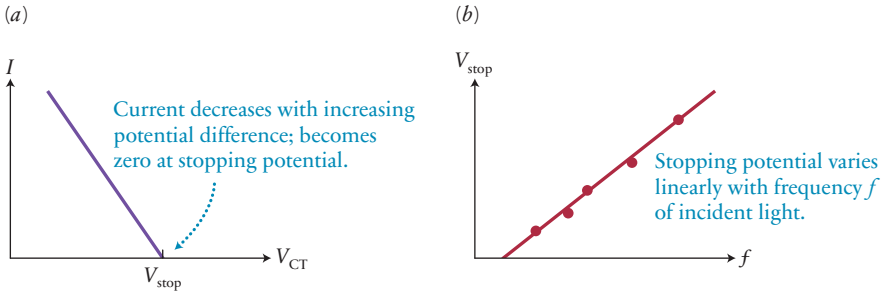


Figure 34.48 For the circuit in Figure 34.47, the current as a function of potential difference and the stopping potential difference as a function of the frequency of the incident light.



shown in the graph of current I versus V_{CT} in **Figure 34.48a**. At this potential difference, called the *stopping potential difference*, the current is zero regardless of the intensity of the incident light. No matter how bright the light, there is no current between the target and the collector. This finding implies that the maximum kinetic energy with which the electrons leave the target does not depend on the intensity (and thus the incident power) of the light.

If we measure the kinetic energy of the ejected electrons (for $V_{CT} < V_{stop}$) directly, we discover that not all of them have the same kinetic energy. This happens because although the amount of energy absorbed from the photons is the same for all electrons, the energy required for the electron to make its way from its initial location in the target to the surface depends on the depth from which the electron is liberated. As a result, electrons released from the surface of the target have the maximum possible kinetic energy, which equals the amount of energy transferred to each electron by the light minus the energy required to liberate the electron from the metal.

For a given potential difference between the target and the collector, the electric field does work $-eV_{CT}$ on an electron as the electron moves from the target to the collector (see Eq. 25.17). The change in the electron's kinetic energy is thus

$$\Delta K = -eV_{CT}. \tag{34.33}$$

Given that the electrons just barely reach the collector at the stopping potential difference, we know that their final kinetic energy is zero, and so for these electrons $\Delta K = K_f - K_i = -K_i$. The maximum kinetic energy with which the electrons leave the target is thus

$$K_{max} = K_i = eV_{stop}. \tag{34.34}$$

Another clue to understanding the experiment in Figure 34.47 emerges when we change the frequency of the incident light and again measure the target-to-collector current as a function of V_{CT}. We observe that the stopping potential difference depends on the frequency; plotting this stopping potential difference as a function of the frequency of the light yields the results shown in Figure 34.48*b*.

34.20 (*a*) What does Figure 34.48*b* tell you about the relationship between the frequency of the incident light and the maximum kinetic energy of the ejected electrons? (*b*) What does the intercept of the line through the data points and the horizontal axis represent?

As Checkpoint 34.20, part *a* shows, the maximum kinetic energy of the ejected electrons depends on the frequency of the incident light, not its intensity. Electrons ejected by ultraviolet light, which has a higher frequency

QUANTITATIVE TOOLS

than visible light, have more kinetic energy than electrons ejected by visible light. (This is why putting glass in the beam of light in the experiment shown in Figure 34.46 essentially eliminates the effect. Although some electrons are ejected by the visible light, those electrons are liberated with less kinetic energy and are more likely to return to the target.) Furthermore, as you discovered in answering Checkpoint 34.20, part *b*, there is a certain minimum frequency of light below which electrons are not ejected at all, regardless of the intensity of the light.

Why does the photoelectric effect require us to think of light as a particle rather than as a wave? The stopping potential difference gives us the maximum kinetic energy with which electrons are released; the light must supply at least this much energy to the electrons in order to eject them. If light could be understood solely as a wave, the intensity of the wave, not its frequency, would determine the maximum amount of energy it could deliver to the electrons. Because the stopping potential difference depends not on light intensity but on frequency, we infer that light carries its energy in energy quanta and that the energy in each quantum is proportional to the frequency.

The photons described in Section 34.5 are these quanta. When an electron absorbs a photon, the electron acquires the photon's entire energy—the electron cannot absorb just part of a photon. The photon's energy frees the electron from the material and gives it additional kinetic energy. If we denote the minimum energy required to free the electron by E_0, we have

$$E_{\text{photon}} = hf = K_{\max} + E_0, \tag{34.35}$$

where $K_{\max}$ is the maximum kinetic energy of the electron as it is ejected. The energy E_0, called the **work function** of the target metal, is a property of the metal that measures how tightly electrons are bound to the metal.

The value of Planck's constant h can be determined by using the relationship between V_{stop} and f given in Figure 34.48*b*. Substituting Eq. 34.34 into Eq. 34.35 and solving the result for V_{stop}, we get

$$V_{\text{stop}} = \frac{h}{e} f - \frac{E_0}{e}. \tag{34.36}$$

This result shows that V_{stop} depends linearly on f and that the slope of the line in Figure 34.48*b* is h/e. By measuring the slope in Figure 34.48*b* and dividing that slope by the charge e of the electron, one obtains $h = 6.626 \times 10^{-34}$ J·s, the value given in Section 34.4.

As discussed in Section 34.5 and expressed in Eq. 34.35, Planck's constant relates the energy and frequency of a photon, $E_{\text{photon}} = hf_{\text{photon}}$, and in Section 34.4 we learned the relationship between the momentum and the wavelength of an electron (or anything else that is ordinarily thought of as a particle): $\lambda_{\text{electron}} = h/p_{\text{electron}}$. Because of the wave-particle duality, we can apply this expression for wavelength to photons as well as electrons: $\lambda_{\text{photon}} = h/p_{\text{photon}}$. If we calculate the momentum of a photon from its wavelength using this expression and then substitute $\lambda = c_0/f$, we see that the momentum of a photon is proportional to its energy:

$$p_{\text{photon}} = \frac{h}{\lambda_{\text{photon}}} = \frac{hf_{\text{photon}}}{c_0} = \frac{E_{\text{photon}}}{c_0}. \tag{34.37}$$

If we substitute this result into the equation relating energy and momentum derived in Chapter 14 (Eq. 14.57)

$$E^2 - (c_0 p)^2 = (mc_0^2)^2, \tag{34.38}$$

the left side of this equation becomes zero, and so we see that photons have zero mass ($m_{\text{photon}} = 0$). We derived Eq. 34.38 for particles that have nonzero mass. Now we see that we can treat photons as massless "particles of light." While these particles have no mass, they do have both momentum and energy:

$$E_{\text{photon}} = hf_{\text{photon}} \tag{34.39}$$

$$p_{\text{photon}} = \frac{hf_{\text{photon}}}{c_0}, \tag{34.40}$$

and, unlike ordinary particles, they always move at the speed of light c_0 in vacuum (or air). Remember also that the mass of a particle is associated with the internal energy of that particle (Eq. 14.54), so $m_{\text{photon}} = 0$ means that photons have no internal energy and therefore no internal structure.

EXAMPLE 34.11 Photoelectric effect

Light of wavelength 380 nm strikes the metal target in Figure 34.47. As long as the potential difference V_{CT} between the target and the collector is no greater than +1.2 V, there is a current in the circuit. Determine the longest wavelength of light that can eject electrons from this metal.

❶ **GETTING STARTED** To solve this problem, I recognize that the longest wavelength of light that can eject electrons corresponds to the lowest-energy photon that can eject an electron; this energy is equal to the work function. I therefore need to use the idea of stopping potential difference to determine the work function from the information given in the problem.

❷ **DEVISE PLAN** The fact that the current is zero when $V_{CT} > +1.2$ V tells me that the stopping potential difference is 1.2 V. Equation 34.36 gives the relationship between photon frequency and stopping potential difference. I can use the relationship between photon frequency and wavelength to rewrite Eq. 34.36 in terms of wavelength. Finally, Eq. 34.35 shows that the lowest photon energy comes when $K_{\max} = 0$, so that the lowest energy equals the work function E_0. Therefore I must determine the wavelength of a photon that has an energy equal to the work function, and for this I can use the expression I developed for the relationship between photon energy and wavelength.

❸ **EXECUTE PLAN** Solving Eq. 34.36 for E_0, then substituting c_0/λ for f and inserting numerical values, I obtain

$$\begin{aligned} E_0 &= \frac{hc_0}{\lambda} - eV_{\text{stop}} \\ &= \frac{(6.626 \times 10^{-34}\ \text{J}\cdot\text{s})(2.998 \times 10^{8}\ \text{m/s})}{380 \times 10^{-9}\ \text{m}} \\ &\quad - (1.602 \times 10^{-19}\ \text{C})(1.2\ \text{V}) \\ &= 3.3 \times 10^{-19}\ \text{J}, \end{aligned}$$

where I have used the equality 1 V ≡ 1 J/C (Eq. 25.16). Because the longest wavelength that can eject electrons has energy equal to the work function, I solve $E_0 = hc_0/\lambda$ for λ and substitute the value of E_0 I just calculated to obtain this maximum wavelength:

$$\begin{aligned} \lambda = \frac{hc_0}{E_0} &= \frac{(6.626\times10^{-34}\ \text{J}\cdot\text{s})(2.998\times10^{8}\ \text{m/s})}{3.3 \times 10^{-19}\ \text{J}} \\ &= 0.60\ \mu\text{m}. \checkmark \end{aligned}$$

❹ **EVALUATE RESULT** This value for the longest wavelength that can eject electrons is greater than 380 nm, the wavelength corresponding to the stopping potential difference of 1.2 V, as it should be.



34.21 A photon enters a piece of glass for which the index of refraction is about 1.5. What happens to the photon's (*a*) speed, (*b*) frequency, (*c*) wavelength, and (*d*) energy?

QUANTITATIVE TOOLS

Chapter Summary

Diffraction (Sections 34.1–34.3, 34.6)

Concepts Light is diffracted as it passes through an opening only when the width of the opening is either comparable to or smaller than the wavelength of the light.

Interference fringes are a pattern of alternating bright and dark bands (called *fringes*) cast on a screen by coherent light passing through two or more slits. **Fringe order** is the numerical ordering of bright and dark fringes. The central bright fringe is the *zeroth-order bright fringe*, the bright fringes to the left and right of it are the first-order bright fringes, the next bright fringes to the left and right are the second-order bright fringes, and so on. The two dark fringes bordering the central bright fringe are the first-order dark fringes, and so on.

When the light passes through two slits, the interference pattern formed on a screen behind the slits consists of alternating bright and dark fringes. When light passes through three or more slits, each dark fringe contains one or more regions that are not as bright as the bright fringes but not completely dark. In these patterns, the brightest fringes are *principal maxima*, and the fainter bright regions are *secondary maxima*.

A **diffraction grating** is a barrier containing a large number of equally spaced slits or grooves. A transmissive diffraction grating contains slits that transmit light, and a reflective diffraction grating contains grooves that reflect light.

The **Bragg condition** states that a crystal reflects x rays of wavelength λ only at angles of incidence relative to the crystal plane given by $2d \cos\theta = m\lambda$, where m is an integer and d is the distance between adjacent crystal planes.

Quantitative Tools When light of wavelength λ passes through two slits a distance d apart, the bright fringes occur at angles given by

$$\sin\theta_m = \pm\frac{m\lambda}{d}, m = 0, 1, 2, 3, \ldots, \tag{34.5}$$

and the dark fringes occur when

$$\sin\theta_n = \pm\frac{\left(n-\frac{1}{2}\right)}{d}\lambda, n = 1, 2, 3, \ldots. \tag{34.7}$$

The integers m and n give the **fringe order.**

For light passing through N slits, the *principal maxima* occur when

$$d\sin\theta_m = \pm m\lambda, m = 0, 1, 2, 3, \ldots, \tag{34.16}$$

and the dark fringes occur when

$$d\sin\theta_{\text{min}} = \pm\frac{k}{N}\lambda, \tag{34.17}$$

where k is an integer that is not an integer multiple of N.

Thin-film interference (Section 34.7)

Concepts Thin-film interference occurs in transparent materials whose thickness is comparable to one of the wavelengths of visible light. Light reflecting from the front surface of the material interferes with light reflecting from the back surface.

Quantitative Tools Light reflects from a thin film of thickness t and index of refraction n_b. At normal incidence of the light, the phase difference ϕ between the light rays reflected from the front and back surfaces of the film is

$$\phi = \frac{4\pi n_b t}{\lambda} + \phi_{r2} - \phi_{r1}, \tag{34.23}$$

where the phase shifts ϕ_{r1} and ϕ_{r2} are due to reflection at the front and back surfaces. Each phase shift must be either 0 or π, depending on the indices of refraction of the media in front of and behind the film.

Diffraction by single slits and circular apertures (Sections 34.8, 34.9)

Concepts Light passing through a circular aperture is diffracted, and the interference fringes are circular. The central bright fringe is called the *Airy disk*.

Quantitative Tools When light is diffracted by a thin slit of thickness a, the dark fringes occur when

$$\sin\theta_n = \pm n\frac{\lambda}{a},\ n = 1, 2, 3, \ldots. \tag{34.26}$$

When light is diffracted by a small circular aperture of diameter d, the first dark fringe occurs when

$$\sin\theta_1 = 1.22\frac{\lambda}{d}. \tag{34.29}$$

Rayleigh's criterion: The minimum angular separation θ_r for which two light sources can be *resolved* by a lens of diameter d is

$$\theta_r \approx 1.22\frac{\lambda}{d}. \tag{34.30}$$

Matter waves and photons (Sections 34.4, 34.5, 34.10)

Concepts **Wave-particle duality** says that particles and light have both wave and particle properties.

The **de Broglie wavelength** expresses the wavelength of a particle.

A **photon** is the "particle" of light. It is an indivisible, discrete unit of light.

Photoelectric effect: When photons of light of sufficiently high frequency shine on a metal, electrons are ejected from the metal. The *stopping potential difference* is the potential needed to stop the ejected electrons. The **work function** E_0 of a metal is the minimum energy needed to free an electron from the metal surface.

Quantitative Tools **Planck's constant** is

$$h = 6.626 \times 10^{-34}\ \text{J}\cdot\text{s}.$$

The **de Broglie wavelength** of a particle having momentum p is

$$\lambda = h/p.$$

In the **photoelectric effect,**

$$E_{\text{photon}} = hf = K_{\max} + E_0, \tag{34.35}$$

where f is the frequency of the light hitting the metal, $K_{\max}$ is the maximum kinetic energy of the ejected electron, and E_0 is the work function of the metal.

The energy of a photon is

$$E_{\text{photon}} = hf_{\text{photon}}, \tag{34.39}$$

and its momentum is

$$p_{\text{photon}} = \frac{hf_{\text{photon}}}{c_0}. \tag{34.40}$$

Questions and Problems

For instructor-assigned homework, go to Mastering Physics

Dots indicate difficulty level of problems: • = *easy*, •• = *intermediate*, ••• = *hard*; CR = *context-rich problem.*

34.1 Diffraction of light

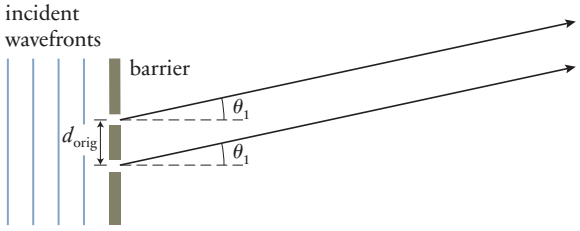
1. Name some characteristics that allow you to observe light diffraction. •
2. Which wavelength of electromagnetic radiation do you expect to be diffracted by a window screen? What is the frequency of this radiation? •
3. Planar waves from a monochromatic light source are normally incident on a circular obstacle, which casts a shadow on a screen positioned behind the obstacle. What do the wave properties of light predict about how dark the center of the shadow is? ••
4. Light of wavelength λ is incident on an aperture of width a, producing diffraction. Describe the change(s) in the diffracted waves (*a*) when the aperture width is doubled and (*b*) when the wavelength is doubled. ••
5. A friend is standing behind a large tree and yelling. You can hear him, but not see him. Why? ••

34.2 Diffraction gratings

6. Coherent green light of wavelength 540 nm passes through two narrow slits for which the center-to-center separation distance is 150 μm. What is the difference in phase between the waves from the two slits for those that arrive at a distant screen at an angular position θ of (*a*) 10.37° and (*b*) 14.6°? •

7. A diffraction grating has adjacent slits separated by 5.00 μm. When yellow light (λ = 570 nm) is incident on the grating, what are the angular positions of the third order bright fringes? •
8. You shine a yellow laser beam on a diffraction grating and then shine a blue laser beam on the same grating. Is the spacing of the bright fringes for the yellow laser beam greater than, smaller than, or equal to the spacing of the bright fringes for the blue laser beam? •
9. A reflection diffraction grating of width w has N grooves and is illuminated by normally incident monochromatic light. What condition determines how many orders of constructive interference can be seen in the interference pattern? ••
10. Light from a He-Cd laser (λ = 442 nm) passing through a two-slit diffraction grating forms a pattern on a screen 2.0 m away from the grating. If the first bright fringe is located at a linear distance y = 22.1 mm to the left of the central bright fringe, what is the slit spacing? What is the position of the second dark fringe? ••
11. Light passing through a two-slit grating makes a pattern on a screen located 500 mm away. In this pattern, the fifth-order dark fringe is 45.0 mm from the central bright fringe. How far from the central bright fringe is the third-order bright fringe? ••
12. A diffraction grating casts a pattern on a screen located a distance L from the grating. The central bright fringe falls directly in the center of the screen. For the highest-order bright fringe that hits the screen, $m = x$, and this fringe hits exactly on the screen edge. This means that $2x + 1$ bright fringes are visible on the screen. What happens to the number of bright fringes on the screen (*a*) if the wavelength of the light passing through the grating is doubled and (*b*) if the spacing d between adjacent slits is doubled? •••
13. (*a*) For the two-slit barrier shown in Figure P34.13, calculate the angle θ_1 for the first-order bright fringe if the slit separation distance is d_{orig} = 1.0 μm and the wavelength of the light passing through the slits is 546 nm. (*b*) Now consider a barrier in which one of the slits oscillates vertically and as a result the slit separation distance varies with time according to $d(t) = d_{orig} + A \sin \omega t$. Calculate θ_1 for the first-order bright fringe as a function of time for A = 0.25 μm and ω = 100 s^{-1}. (*c*) At what instant after t = 0 does θ_1 have its maximum value for the first time? What is this maximum value? (*d*) At what instant after t = 0 does θ_1 have its minimum value for the first time? What is this minimum value? (*e*) Which of the parameters A, ω, λ, d_{orig} affect the values of these angles?

Figure P34.13



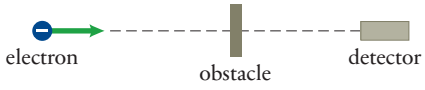
34.3 X-ray diffraction

14. Why are x rays rather than visible light used to determine structure in crystals? •
15. Do the bright spots in an x-ray diffraction pattern get closer together, get farther apart, or not change position when the energy of the x rays is doubled? •
16. A monochromatic x-ray beam that has a wavelength of 0.5×10^{-10} m strikes a sodium chloride crystal. The m = 1 Bragg angle is observed to be 5.12°. What is (*a*) the lattice spacing and (*b*) the highest-order Bragg angle? •
17. In an x-ray diffraction experiment on a crystal that has a cubic lattice, the smallest Bragg angle at which you see a peak in the intensity of diffraction pattern is 15.00°. If the x-ray are produced from a copper source (0.154 nm), what is the atomic spacing in the crystal? Assume that only first-order interference fringes are visible in the diffraction pattern of this crystal. ••
18. You are doing x-ray diffraction on a crystal that has a cubic structure, using 0.1-nm x-rays. If the two greatest Bragg angles at which you observe peaks in the intensity of the diffracted rays are 76° and 63.6°, what is the lattice spacing of the crystal? ••

34.4 Matter waves

19. What is the order of magnitude of the de Broglie wavelength of a person walking down the street? •
20. How fast must a helium atom move in order to have a de Broglie wavelength of 1.5×10^{-10} m? •
21. A narrow solid obstacle is placed in the path of the electron in Figure P34.21. Does the electron have a chance of reaching the detector? •

Figure P34.21



22. What is the de Broglie wavelength of a 0.43-kg soccer ball moving at 30 m/s? •
23. An alpha particle, which consists of two neutrons and two protons, has a mass of 6.645×10^{-27} kg. What is the de Broglie wavelength of an alpha particle that has been accelerated from rest by means of a potential difference of 100 V? ••
24. Rank the following types of radiation and types of particles in order of increasing spacing between bright fringes in their diffraction patterns: yellow light, blue light, electrons moving at a speed of 1.04×10^3 m/s, protons moving at a speed of 1.04×10^3 m/s, and neutrons moving at a speed of 1.04×10^3 m/s. ••
25. You are using a 2000 V potential difference to accelerate electrons toward a diffraction grating and produce a diffraction pattern. If you observe the first intensity peak at an angle of 2.2°, what is the lattice spacing of the crystal? •••

34.5 Photons

26. Which photons contain more energy: those of blue light or those of red light? •
27. In a region of Earth's surface where the average amount of solar energy striking a square meter of ground each second is 200 J, how many photons strike each square meter each second? Use 580 nm for the average wavelength of this solar radiation. •
28. What is the smallest amount of (*a*) momentum and (*b*) energy that can be delivered using violet light for which $\lambda = 400$ nm? •
29. At what rate do photons emitted by a laser ($\lambda = 650$ nm) hit a wall 3.00 m from the laser if the power rating is 6.00 mW? What assumption must you make to solve this problem? ••
30. The sun has a power output of 3.83×10^{26} W. (*a*) How many photons does the Sun emit each second? What assumption(s) must you make to answer this question? (*b*) Installed on your roof is a solar panel that has a surface area of 1.00 m^2. How many photons from the Sun reach the panel? What assumption(s) must you make? ••
31. (*a*) What is the energy of a 540-nm photon emitted by a green laser source? (*b*) There are two green lasers that both emit 540-nm photons. Laser A has a power rating of 1.5 mW, but laser B has a power rating of only 0.75 mW. Why do the power ratings differ when the lasers emit the same wavelength of light? ••

34.6 Multiple-slit interference

32. Near the center of a diffraction pattern, adjacent maxima are separated by a distance of 20 mm. The diffraction grating is 0.200 m from the screen and has a slit spacing of 4.5 μm. What is the wavelength of light being used? •
33. A red laser beam ($\lambda = 647$ nm) is sent through a double slit to a screen that is 15 cm wide and 2.00 m away from the slits. (*a*) If the separation distance between the slits is 200 μm, how many bright fringes appear on the screen? (*b*) Verify your answer to part *a* by calculating the distance between the adjacent bright fringes. •
34. Laser light passes through a double-slit barrier with slit separation $d = 0.120$ mm onto a screen a distance $L = 6.00$ m away. The distance between the central and first-order bright fringes is $y = 27.0$ mm. What is the wavelength and color of the laser light? •
35. Using a diffraction grating with 100 slits/mm to create a diffraction pattern on a screen 2.50 m away from the grating, you see two neighboring bright fringes 150 mm apart. What is the wavelength and the color of the light? •
36. Green light ($\lambda = 510$ nm) shines through two slits separated by 2.30×10^{-6} m. The resulting diffraction pattern is cast on a screen 450 mm away. How far from the central bright fringe is the time-averaged intensity reduced to 12% of its maximum? ••
37. The radiation emitted by mercury vapor lamps includes green 546.1-nm light and blue 435.8-nm light. If this radiation passes through a transmission grating that has 600 slits/mm to form a diffraction pattern on a screen 0.500 m from the grating, what are the positions of the first-order and second-order bright fringes for each wavelength? Report your answers in terms of the linear distance *y* from the central bright fringe. ••
38. Cell phone radiation has a wavelength of about 300 mm. You are in the lobby of an office building in which the façade is a series of large, multistory panes of glass separated by vertical aluminum beams installed every 500 mm across the width of the building. Aluminum shields electromagnetic radiation. You are 10 m from the glass trying to make a phone call, but you don't get any reception. Could this have to do with interference? ••
39. You pass 589.0-nm radiation and 589.6-nm radiation through a diffraction grating that has 500.0 slits/mm. (*a*) In the diffraction pattern, what is the angular separation between the $m = 3$ peak for the 589.0-nm radiation and the $m = 3$ peak for the 589.6-nm radiation? (*b*) Are there fourth-order fringes in the diffraction pattern? ••
40. A beam of 550-nm light illuminates six parallel slits spaced 0.125 mm apart, and the resulting pattern is viewed on a screen. (*a*) What are the angular positions of the first-order and second-order principal maxima? (*b*) What are the angular positions of all the dark fringes between the central maximum and the first-order principal maximum? Between the central maximum and the first-order principal maximum, (*c*) how many dark fringes (minima) are there and (*d*) how many secondary maxima? ••
41. You shine a red ($\lambda = 650$ nm) laser beam and a green ($\lambda = 532$ nm) laser beam at a diffraction grating that has 400 slits/mm to create an interference pattern on a wall 3.00 m from the grating. If the two lasers are directly on top of each other, how far apart are the two $m = 1$ bright fringes from each other? Report your answer as a linear distance. ••
42. Is it possible to shine two laser beams of visible light into the same diffraction grating so that the dark fringes created by one beam overlay the bright fringes created by the other beam for all values of *m*? ••
43. Light of wavelength 570 nm passes through a pair of parallel slits that are 0.115 mm apart. It then falls on a screen that is 45.0 mm from the slits and centered directly opposite them. What must the minimum screen width be in order to have all the bright fringes fall on it? ••
44. A beam of light containing all the visible wavelengths from 400 nm to 700 nm passes through a pair of parallel slits that are 2.50 μm apart. These slits spread the beam into a series of full-color spectra. How many complete full-color spectra are formed? ••
45. You are designing a diffraction grating that, for 566.0-nm light, will produce second-order bright fringes that are an angular distance of 22.00° from the central bright fringe. How many equally spaced slits should this grating contain across its width if it is 24.00 mm wide? ••
46. You want to use a diffraction grating to resolve a beam of visible light into its component wavelengths, from 390 nm to 750 nm. Do the spectra from different orders overlap; that is, does a color from one order overlap with a different color from another order? If so, for which orders? ••
47. The spectrum of the radiation emitted by hydrogen atoms has a peak at $\lambda_r = 656.3$ nm (red) and a peak at $\lambda_b = 486.1$ nm (blue-green). (*a*) If this radiation is passed through a double-slit apparatus in which the slit separation is $d = 0.100$ mm and is then incident on a

screen placed a distance $L = 2.10$ m from the slits, how far apart are the $m = 1$ maxima for the two wavelengths? Can these two maxima be resolved? (*b*) If, instead, a grating with 600 slits/mm is used, how far apart are the maxima now? •••

48. How many slits are needed in a diffraction grating that must resolve the $m = 1$ maxima of two spectral lines at $\lambda_s = 610$ nm and $\lambda_l = 615$ nm? •••

34.7 Thin-film interference

49. A 0.500-μm-thick flake of glass ($n_{glass} = 1.52$) from a broken cover slip is floating on water ($n_{water} = 1.33$). If white light initially traveling in the air ($n_{air} = 1.00$) is incident normal to the surface of the glass, what wavelengths in the visible spectrum undergo constructive interference as the white light is reflected from the top and bottom surfaces of the glass? •
50. You are designing a thin transparent reflective coating for the front surface of a sheet of glass. The index of refraction of the glass is 1.52, and when it is in use the coated glass has air on both sides. Because the coating is expensive, you want to use a layer that has the minimum thickness possible, which you determine to be 104 nm. What should the index of refraction of the coating be if it must cancel 550-nm light that hits the coated surface at normal incidence? ••
51. A sheet of glass ($n_{glass} = 1.5$) is coated with a 90.6-nm-thick layer of magnesium fluoride ($n_{coating} = 1.38$) to prevent reflection in the visible spectrum. What is the longest wavelength of light strongly reflected from this coated surface? Assume the incoming light is normal to the surface. ••
52. The thin film that forms a soap bubble has an index of refraction of 1.42. When white light strikes the outer surface of the film at normal incidence, 625-nm reflected light is especially bright. Derive an expression for the possible thicknesses of the film. ••
53. You are working with the mineral fluorite (CaF_2, $n_{fluorite} = 1.43$) and have a sample that is coated with a layer of liquid 158 nm thick. When you shine visible light of various wavelengths at normal incidence onto the surface of the liquid, you observe very strong reflection for green light ($\lambda = 510$ nm), essentially zero reflection for red light ($\lambda = 750$ nm), and intermediate levels of reflection for all wavelengths between the red and green extremes. What is the index of refraction of the liquid? •••
54. A thin film that is 175 nm thick and has an index of refraction smaller than 1.56 covers the front surface of a vertical sheet of glass that has an index of refraction of 1.56. The back surface of the glass interfaces with air, as does the front surface of the film. When a beam of light of variable wavelength is normally incident on the front surface of the film, we observe that, as the wavelength of the light is varied, 473-nm light reflects strongly from the front film surface. We also determine that if the film is any thinner than 175 nm, no constructive interference occurs. When the beam direction is changed so that the angle of incidence at which the light strikes the air–film interface is 51.0°, what is the angle of refraction for the beam in the film? •••

34.8 Diffraction at a single-slit barrier

55. When a green laser light ($\lambda = 550$ nm) shines on a fracture-line crack in a thin piece of metal, a diffraction pattern is observed on a screen located $L = 0.80$ m from the metal. If the width of the central bright fringe is $w = 10$ mm, what is the width of the crack in the metal? •
56. How wide does a single slit have to be so that 684-nm light passing through the slit has its first dark fringe 20.0° from the center of the interference pattern? •
57. You send light through a slit that is 150 μm wide. What is the wavelength of the light that you must use so that the $n = 1$ fringes are 1.8 cm apart on a screen 2.0 m from the slits? •
58. Blue laser light ($\lambda = 476.5$ nm) passes through a slit of width $a = 0.150$ mm. On a screen a distance $L = 2.500$ m from the slit, what is the linear distance between the central bright fringe and either first-order dark fringe? •
59. Monochromatic 647-nm light is incident on a single slit. If the diffraction pattern is cast on a screen 3.0 m from the slit, and the $n = 3$ dark fringe is 2.5 cm from the center of the pattern, what is the width of the slit? •
60. A student makes the following claim: For radiation of any wavelength passing through a single slit, the separation distance between any two adjacent dark fringes is twice the source–screen distance when the slit width equals the wavelength. Evaluate this claim. ••
61. At sunset, red light travels horizontally through the doorway in the western wall of your beach cabin, and you observe the light on the eastern wall. What is the width of the central maximum in the interference pattern created by the doorway "slit"? ••
62. An experimental setup consists of a 550-nm laser, a screen at a distance of 0.50 m, and an adjustable-width single slit. At what slit width is the width of the central maximum of the interference pattern the same as that of the bright spot it would make in the absence of diffraction? In other words, for what slit width are the effects of wave optics and geometric optics comparable? ••
63. When 440-nm light is incident on a slit that is 75 μm wide, the diffraction pattern is cast on a screen 0.450 m from the slit. How wide is the central bright fringe? ••
64. A slit 0.002470 mm wide is used to study a light ray made up of two wavelengths, 482.0 nm and 517.3 nm, and the diffraction pattern is viewed on a screen 0.220 m from the slit. In the pattern, what is the distance between the $n = 2$ dark fringe for the 482.0-nm light and the $n = 2$ dark fringe for the 517.3-nm light? (Take both dark fringes to be on the same side of the central maximum.) ••
65. A 485-nm light beam passes through a slit and forms a diffraction pattern on a screen 0.320 m from the slit. The $n = 1$ and $n = -1$ dark fringes are 22.4 mm apart on the screen. (*a*) How wide is the slit? (*b*) What is the greatest angle from the original beam direction at which a dark fringe occurs? ••
66. Two extremely narrow, parallel slits are cut in a sheet of cardboard, with the width a of each slit very much smaller than the distance d between them. When the slits are illuminated by a coherent beam of 620-nm light, the angular position of the two third-order bright fringes is $\pm 4.27°$. The cardboard between the slits is then cut out

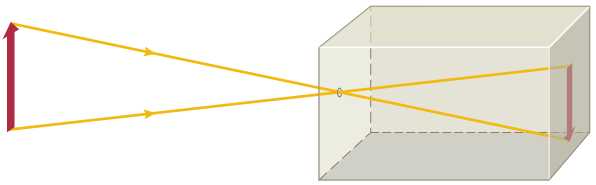
so that the sheet contains only a single opening. When this opening is illuminated by the 620-nm light, what are the angular positions of the $n = 1, 2, 3$ dark fringes? ●●

67. In the interference pattern created by 800-nm radiation passing through a single slit 45 μm wide, what is the angular separation between the $n = 3$ and $n = 5$ dark fringes on the same side of the central maximum? ●●
68. In the interference pattern created by light diffracted from a single slit, which are wider: the first-order bright fringes or the third-order bright fringes? (Hint: Plot or plug in typical numerical values for the slit width and wavelengths to make an educated guess.) ●●●
69. A laser beam passes through a slit that is 1500 nm wide. In the interference pattern created on a distant screen, the angular positions of the two first-order dark fringes are $\pm 25.0°$. The beam then is shone on a soap film ($n_{film} = 1.40$) floating on water ($n_{water} = 1.33$). What is the minimum film thickness that causes light reflected at normal incidence from the air–film interface to cancel light reflected at normal incidence from the film–water interface? ●●●

34.9 Circular apertures and limits of resolution

70. A pinhole of diameter 0.20 mm is illuminated with 550-nm light. What is the width of the Airy disk on a screen 1.5 m away? ●
71. The human eye is most sensitive to green light at 550 nm, which is why this wavelength is most frequently used when calculating the resolution limits of telescopes. A telescope for amateur astronomers has a focal length of 1200 mm and an aperture diameter of 200 mm. What is the radius of the Airy disk? ●
72. You wish to use a lens to focus a beam of light that has a diameter of 40.0 mm. Which lens focuses the beam to the smallest point: lens A, 10.0-mm diameter, 25.0-mm focal length; lens B, 150-mm diameter, 50.0-mm focal length; or lens C, 30.0-mm diameter, 30.0-mm focal length? ●
73. The Spitzer Space Telescope, launched in 2003, has a mirror that is 0.85 m in diameter and detects infrared light with wavelengths from 3.00 μm to 180 μm. If the instrument is used to study a pair of stars, what are the resolvable angular separations of the stars for the minimum and maximum wavelengths? ●
74. For an eye in which the pupil has a radius of 3.0 mm, what is the smallest angular separation that can be resolved (*a*) when two violet ($\lambda = 400$ nm) objects are placed side by side and (*b*) when two red ($\lambda = 650$ nm) objects are placed side by side? When the objects are 100 m away from this eye, what minimum linear separation distance is needed for resolution (*c*) of the two violet objects and (*d*) of the two red objects? ●●
75. A 530-nm laser beam passes through a circular aperture that has diameter 0.400 mm. What is the diameter of the first dark fringe on a screen 800 mm away from the aperture? ●●
76. The pupil of the human eye can vary in diameter from 2.00 mm in bright light to 8.00 mm in dim light. The eye has a focal length of about 25 mm, and the visible spectrum extends from 390 nm (violet) to 750 nm (red). What range of Airy disk radii is possible for the eye? (Note that the light-sensitive cells on the retina have radii ranging from 0.75 μm to 3.0 μm.) ●●
77. In vacuum, the Airy disk generated by a pinhole in a metal sheet has radius $y_{r,vac}$. How does the disk radius change when the sheet is submerged in water ($n = 1.33$)? ●●
78. When 500-nm light is incident on a circular aperture that has diameter $d = 30.0$ μm, a diffraction pattern forms on a screen 350 mm from the aperture. Calculate the area of the central bright fringe. ●●
79. Two objects emitting 550-nm light are placed side by side 30.0 mm apart. For an eye in which the pupil has a diameter of 6.00 mm, what is the minimum eye–objects distance L at which the objects are not resolvable? ●●
80. You are using your telescope to view stars by observing the visible light they emit. If the diameter of the lens is 60.0 mm, what must the minimum angular separation of two stars be in order for you to resolve them? ●●
81. A satellite studying Earth's surface uses a telescope mirror 2.75 m in diameter to focus light of wavelength 525 nm. If the satellite orbits at an altitude of 25,000 km and points the mirror straight downward, what is the diameter of the smallest surface feature it can resolve? ●●
82. A beam of 650-nm light passes through a small round hole and falls on a screen 350 mm past the hole. If the diameter of the Airy disk on the screen is 139 mm, what is the diameter of the hole? ●●
83. The red brake lights of a car are 2.00 m apart. Standing 300 m away from the rear of the car, you use a $f = 50$ mm lens with an aperture of diameter $d = 4.00$ mm to photograph the illuminated brake lights. Are the lights likely to be resolved in the photograph, or will they blur together? ●●●
84. You are building a pinhole camera, which uses a small hole instead of a lens to produce an image (Figure P34.84). (*a*) If the distance between the hole and the film is 100 mm and the hole diameter is 0.300 mm, what is the smallest resolved image point you can have on the film for light ranging from 390 nm to 750 nm? Consider both geometric optics (simply considering the shadow) and wave optics. (*b*) How can you determine the optimum pinhole size for a given wavelength of light? ●●●

Figure P34.84

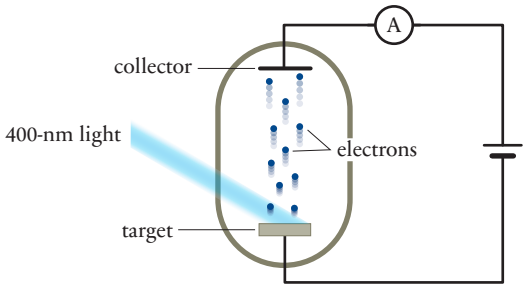


34.10 Photon energy and momentum

85. What happens to (*a*) the kinetic energy of a marble when its speed is reduced to half its initial speed and (*b*) the energy of a photon when its speed is reduced to half its initial speed (as, for instance, when the photon travels from a medium with an index of refraction of 1 into a medium with an index of refraction of 2)? ●
86. What is the difference of the energies of a 500-nm photon and a 600-nm photon? ●
87. What is the wavelength of a gamma-ray photon that has energy of 8.0×10^{-14} J? ●
88. What is the momentum of a photon that has energy of 5.0×10^{-15} J? ●

89. Aluminum has a work function of $E_0 = 6.54 \times 10^{-19}$ J. What is the maximum wavelength of light that can free electrons from the surface? •
90. If the work function of a material is such that red light of wavelength 700 nm just barely initiates the photoelectric effect, what must the maximum kinetic energy of ejected electrons be when violet light of wavelength 400 nm illuminates the material? ••
91. When 420-nm light is incident on a sheet of metal, the electric potential needed to stop the electrons ejected from the metal surface is measured to be 0.67 V. What is the work function of the metal? ••
92. When three metals, 1, 2, and 3, are illuminated with light of frequency f in a photoelectric-effect experiment, the relationship of the stopping potential differences is found to be $V_{\text{stop},1} > V_{\text{stop},2} > V_{\text{stop},3}$. Which material has (*a*) the lowest frequency below which no electrons are ejected and (*b*) the highest frequency below which the effect does not happen? ••
93. In a photoelectric-effect experiment, the stopping potential difference is found to be 3.4 V when 140-nm light is used. What is the work function of the metal? ••
94. A metal alloy has a work function of $E_0 = 4.6 \times 10^{-19}$ J. It is irradiated with light of different wavelengths, and the maximum kinetic energy of ejected electrons is measured. What are the maximum kinetic energy and the maximum electron speed (*a*) when 390-nm light is used and (*b*) when 750-nm light is used? ••
95. You determine that the maximum wavelength of the light needed to eject electrons from the surface of a certain metal is 683 nm. What is the wavelength of the light if the potential needed to stop ejected electrons is 0.8 V? ••
96. A helium-neon laser that has a power rating of 0.250 mW operates at a wavelength of 633 nm and a beam diameter of 2.00 mm. Calculate (*a*) the energy per photon, (*b*) the number of photons emitted per second, (*c*) the momentum per photon, and (*d*) the pressure exerted on an absorber on which the beam is incident. ••
97. (*a*) If the kinetic energy of an electron and the energy of a photon are both 2.00×10^{-18} J, calculate the ratio of the de Broglie wavelength of the electron to the wavelength of the photon. Which particle has the shorter wavelength? (*b*) If the wavelength of a photon and the de Broglie wavelength of an electron are both 250 nm, calculate the ratio of the energy of the photon to the kinetic energy of the electron. Which particle has greater energy? ••
98. A 3.50-W beam of 216-nm laser light shines on the surface of a metal for which the work function is 2.00×10^{-19} J. What is the maximum number of electrons per second the beam can cause to be ejected from the metal? ••
99. A laser beam with an intensity of 60 W/m^2 shines on a black object of mass 2.3 mg. The beam hits an area of 4.5 mm^2 head-on (that is, the beam is perpendicular to the surface). Assuming a totally inelastic collision (meaning the incident photons are absorbed by the black object) and ignoring friction, what is the resulting acceleration of the object? •••
100. In the vacuum tube of Figure P34.100, the lower metal plate, the target, is irradiated with 400-nm light of intensity 5.50 W/m^2. The frequency of the light is great enough to cause electrons to be ejected from the target and travel to the upper metal plate, the collector. What is the maximum possible current in the circuit if the area of the target is 600 mm^2? •••

Figure P34.100



101. When 550-nm light passes through a thin slit and then travels to a screen, the first-order dark fringe in the interference pattern is at 32.5° from the center of the screen. When a beam of electrons, each having kinetic energy equal to the energy of the photons in the 550-nm beam, passes through the slit and travels to the screen, what is the smallest angle from the screen center at which no electrons are found? •••
102. A uniform film of a material that has index of refraction 1.30 covers the front surface of a pane made of glass with index of refraction 1.55. When a beam of monochromatic light initially traveling in air strikes the film at normal incidence, the minimum film thickness for which the light reflected from the air–film interface and the light reflected from the film–glass interface cancel is 123 nm. What is the energy of each photon in the light beam? •••

Additional Problems

103. Monochromatic light of which wavelength diffracts the most through a 3.0-μm aperture: 400 nm, 500 nm, or 600 nm? •
104. A double-slit barrier with slit separation distance d is a distance L from a screen, with $L \gg d$. When green laser light ($\lambda_g = 543$ nm) passes through the barrier, the bright fringes on the screen are a distance y_g apart. When red laser light ($\lambda_r = 694$ nm) passes through the barrier, the bright fringes on the screen are a distance y_r apart. Determine the ratio y_r/y_g. •
105. You cut a slit in a sheet of paper and use it to produce a diffraction pattern on a screen 2 m away from the slit. If, using a laser of wavelength 543.5 nm, you measure a separation distance of 4 mm between the $n = 1$ and $n = 2$ dark fringes of the interference pattern, what is the width of the slit? •
106. NASA plans to use x-ray diffraction to identify minerals on the surface of Mars. If the wavelength of the x rays is 0.155 nm, what is the lattice spacing in a crystalline sample if the crystal strongly reflects x rays at an incident angle of 51.7° in the first order? •
107. Babinet's principle states that, except for differences in intensity, the interference pattern created by light passing around an opaque object is the same as the pattern

created by the same light passing through a hole of the same size and shape as the object. You shine a 690-nm laser beam on a human hair and observe an interference pattern on a screen 1.2 m from the hair. If each first-order dark fringe is 16 mm from the center of the pattern, what is the diameter of the hair? ●●

108. Many hundreds of planets beyond our solar system have been discovered in recent years, but they have all been too far away to be resolved by present-day optical telescopes. Using light of wavelength 525 nm, what must the minimum mirror diameter of a space telescope be in order to resolve an Earth-sized planet at a distance of 10 light years? Are our present space telescopes nearing this size? The diameter of Earth is 1.27×10^7 m. ●●

109. You shine light of frequency f_i on a diffraction grating and create a diffraction pattern on a screen made up of tiny photon detectors. (*a*) How does the pattern change if you increase the frequency of the light? (*b*) How does the pattern change if you replace the initial light source with a source that emits brighter light of the same frequency f_i? (*c*) For the source in part *b*, is the number of photons detected at the screen smaller than, equal to, or greater than the number detected with the initial source? ●●

110. A layer of oil ($n_{oil} = 1.48$) 0.0100 mm thick is resting on a puddle of water ($n_{water} = 1.33$). If white light is incident on the oil, what is the smallest angle from normal at which green light ($\lambda = 510$ nm) is strongly reflected? ●●

111. When a laser beam of wavelength 694 nm passes through a thin slit of width 200 μm, what is the angular separation between the $n = 5$ dark fringe and the adjacent $n = 6$ dark fringe? ●●

112. In a double-slit interference pattern, it is the amplitudes of the light waves from each slit that add, not the light intensities. Usually, in the analysis of interference patterns, the radiation intensity (or amplitude) is assumed to be exactly the same for both slits. Because of inaccurate optical alignment, however, the illumination is often not the same for the two slits. If the intensity of light going through one slit is 0.010 mW/mm^2 while the intensity going through the second slit is 0.030 mW/mm^2, what is the intensity at an interference maximum and at an interference minimum? ●●

113. Electrons are accelerated from rest through a 2.0-kV potential difference. What are (*a*) their speed after this acceleration and (*b*) their wavelength? (*c*) If the accelerated electrons are incident on a graphite crystal that has a lattice spacing of 0.123 nm, do you expect to observe diffraction? ●●

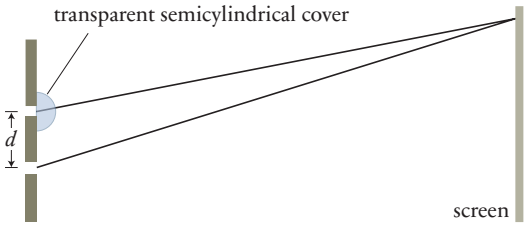
114. You are designing a diffraction grating that will disperse white light into its spectrum (400-nm violet to 700-nm red). In the first-order spectrum, you want the angular separation between the shortest and longest wavelengths to be 12.0°. (*a*) How many slits per millimeter should the grating have? (*b*) How many complete full-color spectra will the grating produce, not counting the central bright fringe? ●●

115. Monochromatic light passes through a small round hole 1.36 μm in radius. The light then strikes a detector 120 mm away from the hole and is absorbed by the detector surface. If the radius of the Airy disk on the detector is 33.3 mm, how much energy and momentum does each photon transfer to the detector? ●●●

116. A young adult with good vision is reading a document placed at her near point. She is using ordinary reading light of wavelength 500 nm, and the diameter of her pupil is 3.0 mm. Under these conditions, what is the height of the smallest letter she can resolve on the document? ●●●

117. In experiment 1, a laser beam of 750-nm light is passed through a double-slit barrier and creates a diffraction pattern on a screen. In experiment 2, one slit is covered with a semicylindrical piece of material that has an index of refraction of $n = 1.001$, as shown in Figure P34.117, and the 750-nm beam again passes through and creates a diffraction pattern on the screen. If the distance between the slits is $d = 1.2$ mm, what is the smallest radius of the semicylindrical piece that results in the screen locations that were bright fringes in experiment 1 being dark fringes in experiment 2, and the screen locations that were dark fringes in experiment 1 being bright fringes in experiment 2? ●●●

Figure P34.117



118. You have an unlabeled container in the laboratory that contains the solid lithium fluoride (LiF) and an identical unlabeled container that contains the solid sodium chloride (NaCl). You know that both materials crystallize in a cubic structure, with alternating Na and Cl ions in the NaCl lattice and alternating Li and F ions in the LiF lattice. The only testing instrument available is a diffraction apparatus that uses monochromatic 0.154-nm x rays. After reading in the literature that the lattice spacing is 0.202 nm for LiF and 0.283 nm for NaCl, you test one of the samples in the apparatus and observe the first intensity peak at a Bragg angle of 22.4°. ●●● CR

119. As a mission engineer for NASA, you are working on a mission to map the surface of Mars. Your boss asks you to design the optical system for a satellite that can resolve surface features as small as 2.00 m across in the visible part of the electromagnetic spectrum. He tells you the instrument has to fit in a cubic bay of 1.0-m width. You know that to get sharp pictures, the satellite must take at least 17 hours to orbit the planet. Luckily you remember your physics classes on gravity, so you sit down to evaluate whether the job is feasible. ●●● CR

Appendix A

Notation

Notation used in this text, listed alphabetically, Greek letters first.
For information concerning superscripts and subscripts, see the explanation at the end of this table.

Symbol	Name of Quantity	Definition	Where Defined	SI units
α (alpha)	polarizability	scalar measure of amount of charge separation occurring in material due to external electric field	Eq. 23.24	$C^2 \cdot m/N$
α	Bragg angle	in x-ray diffraction, angle between incident x rays and sample surface	Section 34.3	degree, radian, or revolution
α_ϑ	(ϑ component of) rotational acceleration	rate at which rotational velocity ω_ϑ increases	Eq. 11.12	s^{-2}
β (beta)	sound intensity level	logarithmic scale for sound intensity, proportional to $\log(I/I_{th})$	Eq. 17.5	dB (not an SI unit)
γ (gamma)	Lorentz factor	factor indicating how much relativistic values deviate from nonrelativistic ones	Eq. 14.6	unitless
γ	surface tension	force per unit length exerted parallel to surface of liquid; energy per unit area required to increase surface area of liquid	Eq. 18.48	N/m
γ	heat capacity ratio	ratio of heat capacity at constant pressure to heat capacity at constant volume	Eq. 20.26	unitless
Δ	delta	change in	Eq. 2.4	
$\Delta\vec{r}$	displacement	vector from object's initial to final position	Eq. 2.8	m
$\Delta\vec{r}_F$, Δx_F	force displacement	displacement of point of application of a force	Eq. 9.7	m
Δt	interval of time	difference between final and initial instants	Table 2.2	s
Δt_{proper}	proper time interval	time interval between two events occurring at same position	Section 14.1	s
Δt_v	interval of time	time interval measured by observer moving at speed v with respect to events	Eq. 14.13	s
Δx	x component of displacement	difference between final and initial positions along x axis	Eq. 2.4	m
δ (delta)	delta	infinitesimally small amount of	Eq. 3.24	
ϵ_0 (epsilon)	electric constant	constant relating units of electrical charge to mechanical units	Eq. 24.7	$C^2/(N \cdot m^2)$
η (eta)	viscosity	measure of fluid's resistance to shear deformation	Eq. 18.38	$Pa \cdot s$
η	efficiency	ratio of work done by heat engine to thermal input of energy	Eq. 21.21	unitless
θ (theta)	angular coordinate	polar coordinate measuring angle between position vector and x axis	Eq. 10.2	degree, radian, or revolution
θ_c	contact angle	angle between solid surface and tangent to liquid surface at meeting point measured within liquid	Section 18.4	degree, radian, or revolution
θ_c	critical angle	angle of incidence greater than which total internal reflection occurs	Eq. 33.9	degree, radian, or revolution
θ_i	angle of incidence	angle between incident ray of light and normal to surface	Section 33.1	degree, radian, or revolution

Symbol	Name of Quantity	Definition	Where Defined	SI units
θ_i	angle subtended by image	angle subtended by image	Section 33.6	degree, radian, or revolution
θ_o	angle subtended by object	angle subtended by object	Section 33.6	degree, radian, or revolution
θ_r	angle of reflection	angle between reflected ray of light and normal to surface	Section 33.1	degree, radian, or revolution
θ_r	minimum resolving angle	smallest angular separation between objects that can be resolved by optical instrument with given aperture	Eq. 34.30	degree, radian, or revolution
ϑ (script theta)	rotational coordinate	for object traveling along circular path, arc length traveled divided by circle radius	Eq. 11.1	unitless
κ (kappa)	torsional constant	ratio of torque required to twist object to rotational displacement	Eq. 15.25	$N \cdot m$
κ	dielectric constant	factor by which potential difference across isolated capacitor is reduced by insertion of dielectric	Eq. 26.9	unitless
λ (lambda)	inertia per unit length	for uniform one-dimensional object, amount of inertia in a given length	Eq. 11.44	kg/m
λ	wavelength	minimum distance over which periodic wave repeats itself	Eq. 16.9	m
λ	linear charge density	amount of charge per unit length	Eq. 23.16	C/m
μ (mu)	reduced mass	product of two interacting objects' inertias divided by their sum	Eq. 6.39	kg
μ	linear mass density	mass per unit length	Eq. 16.25	kg/m
$\vec{\mu}$	magnetic dipole moment	vector pointing along direction of magnetic field of current loop, with magnitude equal to current times area of loop	Section 28.3	$A \cdot m^2$
μ_0	magnetic constant	constant relating units of electric current to mechanical units	Eq. 28.1	$T \cdot m/A$
μ_k	coefficient of kinetic friction	proportionality constant relating magnitudes of force of kinetic friction and normal force between two surfaces	Eq. 10.55	unitless
μ_s	coefficient of static friction	proportionality constant relating magnitudes of force of static friction and normal force between two surfaces	Eq. 10.46	unitless
ρ (rho)	mass density	amount of mass per unit volume	Eq. 1.4	kg/m^3
ρ	inertia per unit volume	for uniform three-dimensional object, amount of inertia in a given volume divided by that volume	Eq. 11.46	kg/m^3
ρ	(volume) charge density	amount of charge per unit volume	Eq. 23.18	C/m^3
σ (sigma)	inertia per unit area	for uniform two-dimensional object, inertia divided by area	Eq. 11.45	kg/m^2
σ	surface charge density	amount of charge per unit area	Eq. 23.17	C/m^2
σ	conductivity	ratio of current density to applied electric field	Eq. 31.8	$A/(V \cdot m)$
τ (tau)	torque	magnitude of axial vector describing ability of forces to change objects' rotational motion	Eq. 12.1	$N \cdot m$
τ	time constant	for damped oscillation, time for energy of oscillator to decrease by factor e^{-1}	Eq. 15.39	s
τ_ϑ	(ϑ component of) torque	ϑ component of axial vector describing ability of forces to change objects' rotational motion	Eq. 12.3	$N \cdot m$
Φ_E (phi, upper case)	electric flux	scalar product of electric field and area through which it passes	Eq. 24.1	$N \cdot m^2/C$

Symbol	Name of Quantity	Definition	Where Defined	SI units
Φ_B	magnetic flux	scalar product of magnetic field and area through which it passes	Eq. 27.10	Wb
ϕ (phi)	phase constant	phase difference between source emf and current in circuit	Eq. 32.16	unitless
$\phi(t)$	phase	time-dependent argument of sine function describing simple harmonic motion	Eq. 15.5	unitless
Ω (omega, upper case)	number of basic states	number of basic states corresponding to macrostate	Section 19.4, Eq. 19.1	unitless
ω (omega)	rotational speed	magnitude of rotational velocity	Eq. 11.7	s^{-1}
ω	angular frequency	for oscillation with period T, $2\pi/T$	Eq. 15.4	s^{-1}
ω_0	resonant angular frequency	angular frequency at which current in circuit is maximal	Eq. 32.47	s^{-1}
ω_ϑ	(ϑ component of) rotational velocity	rate at which rotational coordinate ϑ changes	Eq. 11.6	s^{-1}
A	area	length $\times$ width	Eq. 11.45	m^2
A	amplitude	magnitude of maximum displacement of oscillating object from equilibrium position	Eq. 15.6	m (for linear mechanical oscillation; unitless for rotational oscillation; various units for nonmechanical oscillation)
$\vec{A}$	area vector	vector with magnitude equal to area and direction normal to plane of area	Section 24.6	m^2
$\vec{a}$	acceleration	time rate of change in velocity	Section 3.1	m/s^2
$\vec{a}_{\mathrm{Ao}}$	relative acceleration	value observer in reference frame A records for acceleration of object o in reference frame A	Eq. 6.11	m/s^2
a_c	magnitude of centripetal acceleration	acceleration required to make object follow circular trajectory	Eq. 11.15	m/s^2
a_r	radial component of acceleration	component of acceleration in radial direction	Eq. 11.16	m/s^2
a_t	tangential component of acceleration	component of acceleration tangent to trajectory; for circular motion at constant speed $a_t = 0$	Eq. 11.17	m/s^2
a_x	x component of acceleration	component of acceleration directed along x axis	Eq. 3.21	m/s^2
$\vec{B}$	magnetic field	vector field providing measure of magnetic interactions	Eq. 27.5	T
$\vec{B}_{\mathrm{ind}}$	induced magnetic field	magnetic field produced by induced current	Section 29.4	T
b	damping coefficient	ratio of drag force on moving object to its speed	Eq. 15.34	kg/s
C	heat capacity per particle	ratio of energy transferred thermally per particle to change in temperature	Section 20.3	J/K
C	capacitance	ratio of magnitude of charge on one of a pair of oppositely charged conductors to magnitude of potential difference between them	Eq. 26.1	F
C_P	heat capacity per particle at constant pressure	ratio of energy transferred thermally per particle to change in temperature, while holding pressure constant	Eq. 20.20	J/K
C_V	heat capacity per particle at constant volume	ratio of energy transferred thermally per particle to change in temperature, while holding volume constant	Eq. 20.13	J/K
$\mathrm{COP}_{\mathrm{cooling}}$	coefficient of performance of cooling	ratio of thermal input of energy to work done on a heat pump	Eq. 21.27	unitless

Symbol	Name of Quantity	Definition	Where Defined	SI units
$\mathrm{COP}_{\mathrm{heating}}$	coefficient of performance of heating	ratio of thermal output of energy to work done on a heat pump	Eq. 21.25	unitless
c	shape factor	ratio of object's rotational inertia to mR^2; function of distribution of inertia within object	Table 11.3, Eq. 12.25	unitless
c	wave speed	speed at which mechanical wave travels through medium	Eq. 16.3	m/s
c	specific heat capacity	ratio of energy transferred thermally per unit mass to change in temperature	Section 20.3	J/(K · kg)
c_0	speed of light in vacuum	speed of light in vacuum	Section 14.2	m/s
c_V	specific heat capacity at constant volume	ratio of energy transferred thermally per unit mass to change in temperature, while holding volume constant	Eq. 20.48	J/(K · kg)
$\vec{D}$	displacement (of particle in wave)	displacement of particle from its equilibrium position	Eq. 16.1	m
d	diameter	diameter	Section 1.9	m
d	distance	distance between two locations	Eq. 2.5	m
d	degrees of freedom	number of ways particle can store thermal energy	Eq. 20.4	unitless
d	lens strength	1 m divided by focal length	Eq. 33.22	diopters
E	energy of system	sum of kinetic and internal energies of system	Table 1.1, Eq. 5.21	J
$\vec{E}$	electric field	vector field representing electric force per unit charge	Eq. 23.1	N/C
E_0	work function	minimum energy required to free electron from surface of metal	Eq. 34.35	J
E_{chem}	chemical energy	internal energy associated with object's chemical state	Eq. 5.27	J
E_{int}	internal energy of system	energy associated with an object's state	Eqs. 5.20, 14.54	J
E_{mech}	mechanical energy	sum of system's kinetic and potential energies	Eq. 7.9	J
E_{s}	source energy	incoherent energy used to produce other forms of energy	Eq. 7.7	J
E_{th}	thermal energy	internal energy associated with object's temperature	Eq. 5.27	J
$\mathscr{E}$	emf	in charge-separating device, nonelectrostatic work per unit charge done in separating positive and negative charge carriers	Eq. 26.7	V
$\mathscr{E}_{\mathrm{ind}}$	induced emf	emf resulting from changing magnetic flux	Eqs. 29.3, 29.8	V
$\mathscr{E}_{\mathrm{max}}$	amplitude of emf	amplitude of time-dependent emf produced by AC source	Section 32.1, Eq. 32.1	V
$\mathscr{E}_{\mathrm{rms}}$	rms emf	root-mean-square emf	Eq. 32.55	V
e	coefficient of restitution	measure of amount of initial relative speed recovered after collision	Eq. 5.18	unitless
e	eccentricity	measure of deviation of conic section from circular	Section 13.7	unitless
e	elementary charge	magnitude of charge on electron	Eq. 22.3	C
$\vec{F}$	force	time rate of change of object's momentum	Eq. 8.2	N
$\vec{F}^B$	magnetic force	force exerted on electric current or moving charged particle by magnetic field	Eqs. 27.8, 27.19	N
$\vec{F}^{\mathrm{b}}$	buoyant force	upward force exerted by fluid on submerged object	Eq. 18.12	N
$\vec{F}^{\mathrm{c}}$	contact force	force between objects in physical contact	Section 8.5	N

Symbol	Name of Quantity	Definition	Where Defined	SI units
$\vec{F}^{\mathrm{d}}$	drag force	force exerted by medium on object moving through medium	Eq. 15.34	N
$\vec{F}^{E}$	electric force	force exerted between electrically charged objects or on electrically charged objects by electric field	Eq. 22.1	N
$\vec{F}^{EB}$	electromagnetic force	force exerted on electrically charged objects by electric and magnetic fields	Eq. 27.20	N
$\vec{F}^{\mathrm{f}}$	frictional force	force exerted on object due to friction between it and a second object or surface	Eq. 9.26	N
$\vec{F}^{G}$	gravitational force	force exerted by Earth or any object having mass on any other object having mass	Eqs. 8.16, 13.1	N
$\vec{F}^{k}$	force of kinetic friction	frictional force between two objects in relative motion	Section 10.4, Eq. 10.55	N
$\vec{F}^{n}$	normal force	force directed perpendicular to a surface	Section 10.4, Eq. 10.46	N
$\vec{F}^{s}$	force of static friction	frictional force between two objects not in relative motion	Section 10.4, Eq. 10.46	N
f	frequency	number of cycles per second of periodic motion	Eq. 15.2	Hz
f	focal length	distance from center of lens to focus	Section 33.4, Eq. 33.16	m
f_{beat}	beat frequency	frequency at which beats occur when waves of different frequency interfere	Eq. 17.8	Hz
G	gravitational constant	proportionality constant relating gravitational force between two objects to their masses and separation	Eq. 13.1	$\mathrm{N \cdot m^2/kg^2}$
g	magnitude of acceleration due to gravity	magnitude of acceleration of object in free fall near Earth's surface	Eq. 3.14	$\mathrm{m/s^2}$
h	height	vertical distance	Eq. 10.26	m
h	Planck's constant	constant describing scale of quantum mechanics; relates photon energy to frequency and de Broglie wavelength to momentum of particle	Eq. 34.35	$\mathrm{J \cdot s}$
I	rotational inertia	measure of object's resistance to change in its rotational velocity	Eq. 11.30	$\mathrm{kg \cdot m^2}$
I	intensity	energy delivered by wave per unit time per unit area normal to direction of propagation	Eq. 17.1	$\mathrm{W/m^2}$
I	(electric) current	rate at which charged particles cross a section of a conductor in a given direction	Eq. 27.2	A
I	amplitude of oscillating current	maximum value of oscillating current in circuit	Section 32.1, Eq. 32.5	A
I_{cm}	rotational inertia about center of mass	object's rotational inertia about an axis through its center of mass	Eq. 11.48	$\mathrm{kg \cdot m^2}$
I_{disp}	displacement current	current-like quantity in Ampère's law caused by changing electric flux	Eq. 30.7	A
I_{enc}	enclosed current	current enclosed by Ampèrian path	Eq. 28.1	A
I_{ind}	induced current	current in loop caused by changing magnetic flux through loop	Eq. 29.4	A
I_{int}	intercepted current	current intercepted by surface spanning Ampèrian path	Eq. 30.6	A
I_{rms}	rms current	root-mean-square current	Eq. 32.53	A
I_{th}	intensity at threshold of hearing	minimum intensity audible to human ear	Eq. 17.4	$\mathrm{W/m^2}$
i	time-dependent current	time-dependent current through circuit; $I(t)$	Section 32.1, Eq. 32.5	A
i	image distance	distance from lens to image	Section 33.6, Eq. 33.16	m

Symbol	Name of Quantity	Definition	Where Defined	SI units
$\hat{i}$	unit vector ("i hat")	vector for defining direction of x axis	Eq. 2.1	unitless
$\vec{J}$	impulse	amount of momentum transferred from environment to system	Eq. 4.18	$kg \cdot m/s$
$\vec{J}$	current density	current per unit area	Eq. 31.6	A/m^2
J_ϑ	rotational impulse	amount of angular momentum transferred from environment to system	Eq. 12.15	$kg \cdot m^2/s$
$\hat{j}$	unit vector ("j hat")	vector for defining direction of y axis	Eq. 10.4	unitless
K	kinetic energy	energy object has because of its translational motion	Eqs. 5.12, 14.51	J
K	surface current density	current per unit of sheet width	Section 28.5	A/m
K_{cm}	translational kinetic energy	kinetic energy associated with motion of center of mass of system	Eq. 6.32	J
K_{conv}	convertible kinetic energy	kinetic energy that can be converted to internal energy without changing system's momentum	Eq. 6.33	J
K_{rot}	rotational kinetic energy	energy object has due to its rotational motion	Eq. 11.31	J
k	spring constant	ratio of force exerted on spring to displacement of free end of spring	Eq. 8.18	N/m
k	wave number	number of wavelengths in 2π units of distance; for wave with wavelength λ, $2\pi/\lambda$	Eqs. 16.7, 16.11	m^{-1}
k	Coulomb's law constant	constant relating electrostatic force to charges and their separation distance	Eq. 22.5	$N \cdot m^2/C^2$
k_B	Boltzmann constant	constant relating thermal energy to absolute temperature	Eq. 19.39	J/K
L	inductance	negative of ratio of induced emf around loop to rate of change of current in loop	Eq. 29.19	H
L_ϑ	(ϑ component of) angular momentum	capacity of object to make other objects rotate	Eq. 11.34	$kg \cdot m^2/s$
L_m	specific transformation energy for melting	energy transferred thermally per unit mass required to melt substance	Eq. 20.55	J/kg
L_v	specific transformation energy for vaporization	energy transferred thermally per unit mass required to vaporize substance	Eq. 20.55	J/kg
ℓ	length	distance or extent in space	Table 1.1	m
ℓ_{proper}	proper length	length measured by observer at rest relative to object	Section 14.3	m
ℓ_v	length	measured length of object moving at speed v relative to observer	Eq. 14.28	m
M	magnification	ratio of signed image height to object height	Eq. 33.17	unitless
M_θ	angular magnification	ratio of angle subtended by image to angle subtended by object	Eq. 33.18	unitless
m	mass	amount of substance	Table 1.1, Eq. 13.1	kg
m	inertia	measure of object's resistance to change in its velocity	Eq. 4.2	kg
m	fringe order	number indexing bright interference fringes, counting from central, zeroth-order bright fringe	Section 34.2, Eq. 34.5	unitless
m_v	inertia	inertia of object moving at speed v relative to observer	Eq. 14.41	kg
N	number of objects	number of objects in sample	Eq. 1.3	unitless
N_A	Avogadro's number	number of particles in 1 mol of a substance	Eq. 1.2	unitless
n	number density	number of objects per unit volume	Eq. 1.3	m^{-3}

Symbol	Name of Quantity	Definition	Where Defined	SI units
n	windings per unit length	in a solenoid, number of windings per unit length	Eq. 28.4	unitless
n	index of refraction	ratio of speed of light in vacuum to speed of light in a medium	Eq. 33.1	unitless
n	fringe order	number indexing dark interference fringes, counting from central, zeroth-order bright fringe	Section 34.2, Eq. 34.7	unitless
O	origin	origin of coordinate system	Section 10.2	
o	object distance	distance from lens to object	Section 33.6, Eq. 33.16	m
P	power	time rate at which energy is transferred or converted	Eq. 9.30	W
P	pressure	force per unit area exerted by fluid	Eq. 18.1	Pa
P_{atm}	atmospheric pressure	average pressure in Earth's atmosphere at sea level	Eq. 18.3	Pa
P_{gauge}	gauge pressure	pressure measured as difference between absolute pressure and atmospheric pressure	Eq. 18.16	Pa
p	time-dependent power	time-dependent rate at which source delivers energy to load; $P(t)$	Eq. 32.49	W
$\vec{p}$	momentum	vector that is product of an object's inertia and velocity	Eq. 4.6	kg · m/s
$\vec{p}$	(electric) dipole moment	vector representing magnitude and direction of electric dipole, equal amounts of positive and negative charge separated by small distance	Eq. 23.9	C · m
$\vec{p}_{\text{ind}}$	induced dipole moment	dipole moment induced in material by external electric field	Eq. 23.24	C · m
p_x	x component of momentum	x component of momentum	Eq. 4.7	kg · m/s
Q	quality factor	for damped oscillation, number of cycles for energy of oscillator to decrease by factor $e^{-2\pi}$	Eq. 15.41	unitless
Q	volume flow rate	rate at which volume of fluid crosses section of tube	Eq. 18.25	m^3/s
Q	energy transferred thermally	energy transferred into system by thermal interactions	Eq. 20.1	J
Q_{in}	thermal input of energy	positive amount of energy transferred into system by thermal interactions	Sections 21.1, 21.5	J
Q_{out}	thermal output of energy	positive amount of energy transferred out of system by thermal interactions	Sections 21.1, 21.5	J
q	electrical charge	attribute responsible for electromagnetic interactions	Eq. 22.1	C
q_{enc}	enclosed charge	sum of all charge within a closed surface	Eq. 24.8	C
q_{p}	dipole charge	charge of positively charged pole of dipole	Section 23.6	C
R	radius	radius of an object	Eq. 11.47	m
R	resistance	ratio of applied potential difference to resulting current	Eqs. 29.4, 31.10	Ω
R_{eq}	equivalent resistance	resistance that could be used to replace combination of circuit elements	Eqs. 31.26, 31.33	Ω
r	radial coordinate	polar coordinate measuring distance from origin of coordinate system	Eq. 10.1	m
$\vec{r}$	position	vector for determining position	Eqs. 2.9, 10.4	m
$\hat{r}_{12}$	unit vector ("r hat")	unit vector pointing from tip of $\vec{r}_1$ to tip of $\vec{r}_2$	Eq. 22.6	unitless
$\vec{r}_{\text{AB}}$	relative position	position of observer B in reference frame of observer A	Eq. 6.3	m
$\vec{r}_{\text{Ae}}$	relative position	value observer in reference frame A records for position at which event e occurs	Eq. 6.3	m

Symbol	Name of Quantity	Definition	Where Defined	SI units
$\vec{r}_{\text{cm}}$	position of a system's center of mass	a fixed position in a system that is independent of choice of reference frame	Eq. 6.24	m
$\vec{r}_{\text{p}}$	dipole separation	position of positively charged particle relative to negatively charged particle in dipole	Section 23.6	m
$r_{\perp}$	lever arm distance *or* lever arm	perpendicular distance between rotation axis and line of action of a vector	Eq. 11.36	m
$\Delta\vec{r}$	displacement	vector from object's initial to final position	Eq. 2.8	m
$\Delta\vec{r}_F$	force displacement	displacement of point of application of a force	Eq. 9.7	m
S	entropy	logarithm of number of basic states	Eq. 19.4	unitless
S	intensity	intensity of electromagnetic wave	Eq. 30.36	W/m^2
$\vec{S}$	Poynting vector	vector representing flow of energy in combined electric and magnetic fields	Eq. 30.37	W/m^2
s	arc length	distance along circular path	Eq. 11.1	m
s^2	space-time interval	invariant measure of separation of events in space-time	Eq. 14.18	m^2
T	period	time interval needed for object in circular motion to complete one revolution	Eq. 11.20	s
T	absolute temperature	quantity related to rate of change of entropy with respect to thermal energy	Eq. 19.38	K
$\mathcal{T}$	tension	stress in object subject to opposing forces stretching the object	Section 8.6	N
t	instant in time	physical quantity that allows us to determine the sequence of related events	Table 1.1	s
t_{Ae}	instant in time	value observer A measures for instant at which event e occurs	Eq. 6.1	s
Δt	interval of time	difference between final and initial instants	Table 2.2	s
Δt_{proper}	proper time interval	time interval between two events occurring at same position	Section 14.1	s
Δt_v	interval of time	time interval between two events measured by observer moving at speed v relative to an observer for whom the events occur at the same position	Eq. 14.13	s
U	potential energy	energy stored in reversible changes to system's configuration state	Eq. 7.7	J
U^B	magnetic potential energy	potential energy stored in magnetic field	Eqs. 29.25, 29.30	J
U^E	electric potential energy	potential energy due to relative position of charged objects	Eq. 25.8	J
U^G	gravitational potential energy	potential energy due to relative position of gravitationally interacting objects	Eqs. 7.13, 13.14	J
u_B	energy density of magnetic field	energy per unit volume stored in magnetic field	Eq. 29.29	J/m^3
u_E	energy density of electric field	energy per unit volume stored in electric field	Eq. 26.6	J/m^3
V	volume	amount of space occupied by an object	Table 1.1	m^3
V_{AB}	potential difference	negative of electrostatic work per unit charge done on charged particle as it is moved from point A to point B	Eq. 25.15	V
V_{batt}	battery potential difference	magnitude of potential difference between terminals of battery	Eq. 25.19	V
V_C	amplitude of oscillating potential	maximum magnitude of potential across circuit element C	Section 32.1, Eq. 32.8	V
V_{disp}	displaced volume	volume of fluid displaced by submerged object	Eq. 18.12	m^3

Symbol	Name of Quantity	Definition	Where Defined	SI units
V_P	(electrostatic) potential	potential difference between conveniently chosen reference point of potential zero and point P	Eq. 25.30	V
V_{rms}	rms potential	root-mean-square potential difference	Eq. 32.55	V
V_{stop}	stopping potential	minimum potential difference required to stop flow of electrons from photoelectric effect	Eq. 34.34	V
$\mathcal{V}$	"volume" in velocity space	measure of range of velocities in three dimensions	Eq. 19.20	$(m/s)^3$
v	speed	magnitude of velocity	Table 1.1	m/s
$\vec{v}$	velocity	time rate of change in position	Eq. 2.23	m/s
$\vec{v}_{12}$	relative velocity	velocity of object 2 relative to object 1	Eq. 5.1	m/s
$\vec{v}_{AB}$	relative velocity	velocity of observer B in reference frame of observer A	Eq. 6.3	m/s
v_C	time-dependent potential	time-dependent potential across circuit element C; $V_C(t)$	Section 32.1, Eq. 32.8	V
$\vec{v}_{cm}$	velocity, center of mass	velocity of the center of mass of a system, equal to the velocity of the zero-momentum reference frame of the system	Eq. 6.26	m/s
$\vec{v}_d$	drift velocity	average velocity of electrons in conductor in presence of electric field	Eq. 31.3	m/s
v_{esc}	escape speed	minimum launch speed required for object to reach infinity	Eq. 13.23	m/s
v_r	radial component of velocity	for object moving along circular path, always zero	Eq. 11.18	m/s
v_{rms}	root-mean-square speed	square root of average of square of speed	Eq. 19.21	m/s
v_t	tangential component of velocity	for object in circular motion, rate at which arc length is swept out	Eq. 11.9	m/s
v_x	x component of velocity	component of velocity directed along x axis	Eq. 2.21	m/s
W	work	change in system's energy due to external forces exerted on system	Eqs. 9.1, 10.35	J
$W_{P\to Q}$	work	work done along path from P to Q	Eq. 13.12	J
W_{in}	mechanical input of energy	positive amount of mechanical work done on system	Section 21.1	J
W_{out}	mechanical output of energy	positive amount of mechanical work done by system	Section 21.1	J
W_q	electrostatic work	work done by electrostatic field on charged particle moving through field	Section 25.2, Eq. 25.17	J
X_C	capacitive reactance	ratio of potential difference amplitude to current amplitude for capacitor	Eq. 32.14	Ω
X_L	inductive reactance	ratio of potential difference amplitude to current amplitude for inductor	Eq. 32.26	Ω
x	position	position along x axis	Eq. 2.4	m
$x(t)$	position as function of time	position x at instant t	Section 2.3	m
Δx	x component of displacement	difference between final and initial positions along x axis	Eq. 2.4	m
Δx_F	force displacement	displacement of point of application of a force	Eq. 9.7	m
Z	impedance	(frequency-dependent) ratio of potential difference to current through circuit	Eq. 32.33	Ω
z	zero-momentum reference frame	reference frame in which system of interest has zero momentum	Eq. 6.23	

Math notation

Math notation	Name	Where introduced
$\equiv$	defined as	Eq. 1.3
$\approx$	approximately equal to	Section 1.9
Σ (sigma, upper case)	sum of	Eq. 3.25
$\int$	integral of	Eq. 3.27
$\parallel$	parallel	Section 10.2
$\perp$	perpendicular	Section 10.2
$\propto$	proportional to	Section 13.1
$\cdot$	scalar product of two vectors	Eq. 10.33
$\times$	vector product of two vectors	Eq. 12.35
$\frac{\partial f}{\partial x}$	partial derivative of f with respect to x	Eq. 16.47
$\vec{b}$	vector b	Eq. 2.2
$\lvert\vec{b}\rvert$ or b	magnitude of $\vec{b}$	Eq. 2.3
b_x	x component of $\vec{b}$	Eq. 2.2
$\vec{b}_x$	x component vector of $\vec{b}$	Eq. 10.5
$\hat{\imath}$	unit vector ("i hat")	Eq. 2.1
$\hat{r}_{12}$	unit vector ("r hat")	Eq. 22.6

Note concerning superscripts and subscripts

Superscripts are appended to forces and potential energies to indicate the type of force or energy. They may be found in the main list under F, for forces, and U, for potential energies. Uppercase superscripts are used for fundamental interactions.

Subscripts are used on many symbols to identify objects, reference frames, types (for example, of energy), and processes. Object identifiers may be numbers, letters, or groups of letters. Reference frames are indicated by capital letters. Object identifiers and reference frames can occur in pairs, indicating relative quantities. In this case, the main symbol describes a property of whatever is identified by the second subscript relative to that of the first. In the case of forces, the first subscript identifies the object that causes the force and the second identifies the object on which the force is exerted. Types and processes are identified in various ways; many are given in the main list. Here are some examples:

m_1	inertia of object 1
m_{ball}	inertia of ball
$\vec{v}_{\text{cm}}$	velocity of center of mass of system
$\vec{r}_{12}$	position of object 2 relative to object 1; $\vec{r}_{12} = \vec{r}_2 - \vec{r}_1$
$\vec{p}_1$	momentum of object 1
$\vec{p}_{\text{Z2}}$	momentum of object 2 as measured in zero-momentum reference frame
$\vec{v}_{\text{AB}}$	velocity of observer B as measured in reference frame of observer A
$\vec{v}_{\text{Ao}}$	velocity of object o as measured in reference frame A
$\vec{r}_{\text{Ee}}$	position of event e as measured in Earth reference frame
$\vec{F}^{\text{c}}_{\text{pw}}$	contact force exerted by person on wall
$\vec{F}^{G}_{\text{Eb}}$	gravitational force exerted by Earth on ball
E_{th}	thermal energy
K_{conv}	convertible kinetic energy
P_{av}	average power
a_{c}	centripetal acceleration
$W_{\text{P}\to\text{Q}}$	work done along path from P to Q

Initial and final conditions are identified by subscripts i and f, following other identifiers. For example:

$\vec{p}_{1\text{i}}$	initial momentum of object 1
$\vec{p}_{\text{Zball,f}}$	final momentum of ball as measured in zero-momentum reference frame

Italic subscripts are used to identify components of vectors. These include x, y, z, r (radial), t (tangential), and ϑ (angular, with respect to given axis). They are also used to enumerate collections, for example, as indices of summation, and to indicate that a subscript refers to another variable. Here are some examples:

r_x	x component of position
a_t	tangential component of acceleration
L_ϑ	ϑ component of angular momentum
$p_{\text{Zball}\,y,\text{f}}$	final y component of momentum of ball as measured in zero-momentum reference frame
$\delta m_n r_n^2$	contribution to rotational inertia of extended object of small segment n, with inertia δm_n, at position r_n
c_P	specific heat capacity at constant pressure
W_q	electrostatic work

Appendix B

Mathematics Review

1 Algebra

Factors

$$ax + bx + cx = (a + b + c)x$$
$$(a + b)^2 = a^2 + 2ab + b^2$$
$$(a - b)^2 = a^2 - 2ab + b^2$$
$$(a + b)(a - b) = a^2 - b^2$$

Fractions

$$\left(\frac{a}{b}\right)\left(\frac{c}{d}\right) = \frac{ac}{bd}$$
$$\left(\frac{a/b}{c/d}\right) = \frac{a}{b} \div \frac{c}{d} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}$$
$$\left(\frac{1}{1/a}\right) = a$$

Exponents

$$a^n = \underbrace{a \times a \times a \times \cdots \times a}_{n \text{ factors}}$$

Any real number can be used as an exponent:

$$a^{-x} = \frac{1}{a^x}$$
$$a^0 = 1$$
$$a^1 = a$$
$$a^{1/2} = \sqrt{a}$$
$$a^{1/n} = \sqrt[n]{a}$$
$$a^x a^y = a^{x+y}$$
$$\frac{a^x}{a^y} = a^{x-y}$$
$$(a^x)^y = a^{x \cdot y}$$
$$a^x b^x = (ab)^x$$
$$\frac{a^x}{b^x} = \left(\frac{a}{b}\right)^x$$

Logarithms

Logarithm is the inverse function of the exponential function:

$$y = a^x \Leftrightarrow \log_a y = \log_a a^x = x \quad \text{and} \quad x = \log_a (a^x) = a^{\log_a x}$$

The two most common values for the base a are 10 (the common logarithm base) and e (the natural logarithm base).

$$y = e^x \Leftrightarrow \log_e y = \ln y = \ln e^x = x \quad \text{and} \quad x = \ln e^x = e^{\ln x}$$

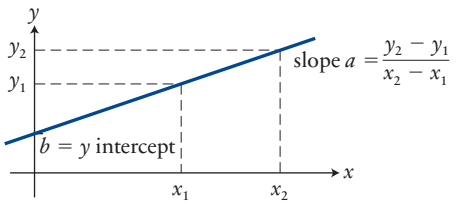
Logarithm rules (valid for any base):

$$\ln(ab) = \ln(a) + \ln(b)$$
$$\ln\left(\frac{a}{b}\right) = \ln(a) - \ln(b)$$
$$\ln(a^n) = n \ln(a)$$
$$\ln 1 = 0$$

The expression $\ln(a + b)$ cannot be simplified.

Linear equations

A linear equation has the form $y = ax + b$, where a and b are constants. A graph of y versus x is a straight line. The value of a equals the slope of the line, and the value of b equals the value of y when x equals zero.



If $a = 0$, the line is horizontal. If $a > 0$, the line rises as x increases. If $a < 0$, the line falls as x increases. For any two values of x, say x_1 and x_2, the slope a can be calculated as

$$a = \frac{y_2 - y_1}{x_2 - x_1}$$

where y_1 and y_2 correspond to x_1 and x_2 (that is to say, $y_1 = ax_1 + b$ and $y_2 = ax_2 + b$).

Proportionality

If y is proportional to x (written $y \propto x$), then $y = ax$, where a is a constant. Proportionality is a subset of linearity. Because $y/x = a =$ constant for any corresponding x and y,

$$\frac{y_1}{x_1} = \frac{y_2}{x_2} \Leftrightarrow \frac{y_1}{y_2} = \frac{x_1}{x_2}.$$

Quadratic equation

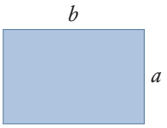
The equation $ax^2 + bx + c = 0$ (the quadratic equation) has two solutions (called *roots)* for x:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

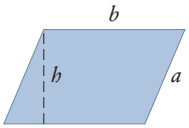
If $b^2 \geq 4ac$, the solutions are real numbers.

2 Geometry

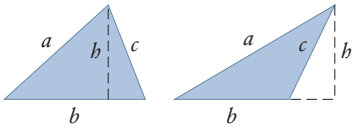
Area and circumference for two-dimensional shapes



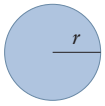
rectangle:
area $= ab$
circumference $= 2(a + b)$



parallelogram:
area $= bh$
circumference $= 2(a + b)$

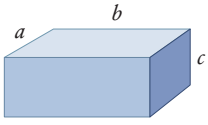


triangle:
area $= \frac{1}{2}bh$
circumference $= a + b + c$

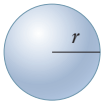


circle:
area $= \pi r^2$
circumference $= 2\pi r$

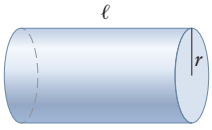
Volume and area for three-dimensional shapes



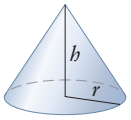
rectangular box:
volume $= abc$
area $= 2(a^2 + b^2 + c^2)$



sphere:
volume $= \frac{4}{3}\pi r^3$
area $= 4\pi r^2$



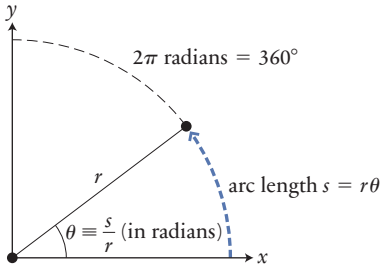
right circular cylinder:
volume $= \pi r^2 \ell$
area $= 2\pi r\ell + 2\pi r^2$



right circular cone:
volume $= \frac{1}{3}\pi r^2 h$
area $= \pi r^2 + \pi r\sqrt{r^2 + h^2}$

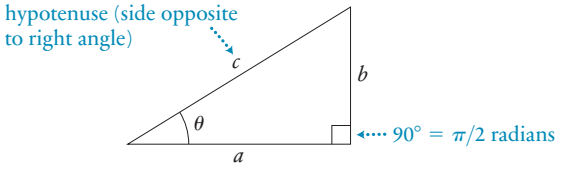
3 Trigonometry

Angle and arc length



Right triangles

A right triangle is a triangle in which one of the angles is a right angle:



Pythagorean theorem: $a^2 + b^2 = c^2 \Leftrightarrow c = \sqrt{a^2 + b^2}$
Trigonometric functions:

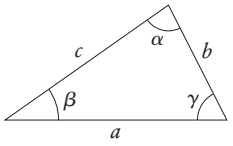
$$\sin\theta = \frac{b}{c} = \frac{\text{opposite side}}{\text{hypotenuse}}, \quad \theta = \sin^{-1}\left(\frac{b}{c}\right) = \arcsin\left(\frac{b}{c}\right)$$

$$\cos\theta = \frac{a}{c} = \frac{\text{adjacent side}}{\text{hypotenuse}}, \quad \theta = \cos^{-1}\left(\frac{a}{c}\right) = \arccos\left(\frac{a}{c}\right)$$

$$\tan\theta = \frac{b}{a} = \frac{\text{opposite side}}{\text{adjacent side}}, \quad \theta = \tan^{-1}\left(\frac{b}{a}\right) = \arctan\left(\frac{b}{a}\right)$$

General triangles

For any triangle, the following relationships hold:



$$\alpha + \beta + \gamma = 180° = \pi \text{ rad}$$

$$\text{Sine law: } \frac{\sin\alpha}{a} = \frac{\sin\beta}{b} = \frac{\sin\gamma}{c}$$

$$\text{Cosine law: } c^2 = a^2 + b^2 - 2ab\cos\gamma$$

Identities

$$\tan\theta = \frac{\sin\theta}{\cos\theta}$$

$$\cot\theta = \frac{1}{\tan\theta} = \frac{\cos\theta}{\sin\theta}$$

$$\csc\theta = \frac{1}{\sin\theta}$$

$$\sec\theta = \frac{1}{\cos\theta}$$

Periodicity

$$\cos(\alpha + 2\pi) = \cos\alpha$$
$$\tan(\alpha + \pi) = \sin\alpha$$

Angle addition

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta$$
$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$$

Double angles

$$\sin(2\alpha) = 2\sin\alpha\cos\alpha$$

$$\cos(2\alpha) = \cos^2\alpha - \sin^2\alpha = 1 - 2\sin^2\alpha = 2\cos^2\alpha - 1$$

Other relations

$$\sin^2\alpha + \cos^2\alpha = 1$$

$$\sin(-\alpha) = -\sin\alpha$$

$$\cos(-\alpha) = \cos\alpha$$

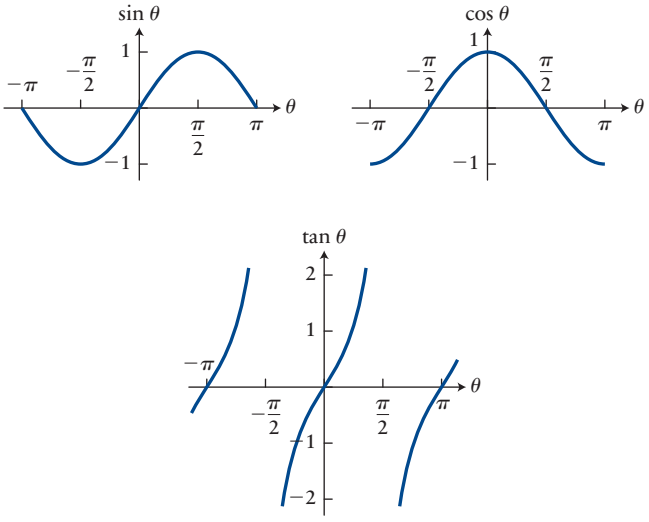
$$\sin(\alpha \pm \pi) = -\sin\alpha$$

$$\cos(\alpha \pm \pi) = -\cos\alpha$$

$$\sin(\alpha \pm \pi/2) = \pm\cos\alpha$$

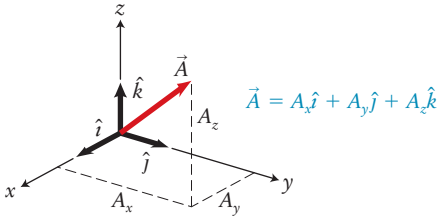
$$\cos(\alpha \pm \pi/2) = \mp\sin\alpha$$

The following graphs show sin θ, cos θ, and tan θ as functions of θ:



4 Vector algebra

A vector $\vec{A}$ in three-dimensional space can be written in terms of magnitudes A_x, A_y, and A_z of unit vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$, which have length 1 and lie along the x, y, and z axes:



Dot products between vectors produce scalars:

$$\vec{A}\cdot\vec{B} = A_xB_x + A_yB_y + A_zB_z = |A||B|\cos\theta$$

(θ is the angle between vectors $\vec{A}$ and $\vec{B}$)

Cross products between vectors produce vectors:

$$\vec{A}\times\vec{B} = (A_yB_z - A_zB_y)\hat{\imath} + (A_zB_x - A_xB_z)\hat{\jmath} + (A_xB_y - A_yB_x)\hat{k}$$

$$|\vec{A}\times\vec{B}| = |\vec{A}||\vec{B}|\sin\theta$$ (θ is the angle between vectors $\vec{A}$ and $\vec{B}$)

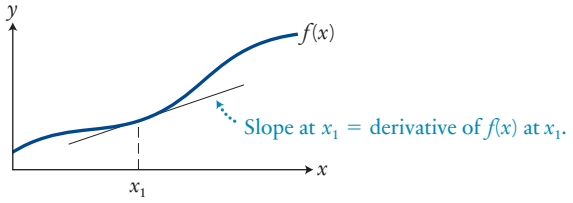
The direction of $\vec{A}\times\vec{B}$ is given by the right-hand rule (see Figure 12.44).

5 Calculus

In this section, x is a variable, and a and n are constants.

Derivatives

Geometrically, the derivative of a function $f(x)$ at $x = x_1$ is the slope of $f(x)$ at x_1:



Derivatives of common functions

$$\frac{d}{dx}a = 0$$

$$\frac{d}{dx}x^n = nx^{n-1}$$ (n need not be an integer)

$$\frac{d}{dx}\sin x = \cos x$$

$$\frac{d}{dx}\cos x = -\sin x$$

$$\frac{d}{dx}\tan x = \frac{1}{\cos^2 x}$$

$$\frac{d}{dx}e^{ax} = ae^{ax}$$

$$\frac{d}{dx}\ln(ax) = \frac{1}{x}$$

$$\frac{d}{dx}a^x = a^x\ln a$$

Derivatives of sums, products, and functions of functions

Constant times a function: $\frac{d}{dx}[a\cdot f(x)] = a\cdot\frac{d}{dx}f(x)$

Sum of functions: $\frac{d}{dx}[f(x) + g(x)] = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$

Product of functions:

$$\frac{d}{dx}[f(x)\cdot g(x)] = g(x)\frac{d}{dx}f(x) + f(x)\frac{d}{dx}g(x)$$

Quotient of functions: $\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)\frac{d}{dx}f(x) - f(x)\frac{d}{dx}g(x)}{[g(x)]^2}$

Functions of functions (the chain rule): If f is a function of u, and u is a function of x, then

$$\frac{d[f(u)]}{du}\cdot\frac{d[u(x)]}{dx}=\frac{d[f(x)]}{dx}$$

Second and higher derivatives The second derivative of a function f with respect to x is the derivative of the derivative:

$$\frac{d^2f(x)}{dx^2}=\frac{d}{dx}\left(\frac{d}{dx}f(x)\right)$$

Higher derivatives are defined similarly:

$$\frac{d^nf(x)}{dx^n}=\underbrace{\cdots\frac{d}{dx}\left(\frac{d}{dx}\left(\frac{d}{dx}f(x)\right)\right)}_{n\text{ uses of }\frac{d}{dx}}\quad\text{(where }n\text{ is a positive integer).}$$

Partial derivatives For functions of more than one variable, the partial derivative, written $\frac{\partial}{\partial x}$, is the derivative with respect to one variable; all other variables are treated as constants.

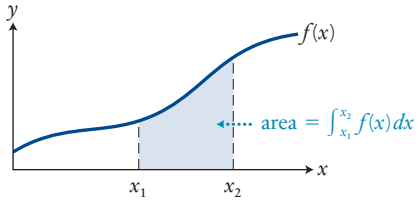
Integrals

Indefinite integrals Integration is the reverse of differentiation. An indefinite integral $\int f(x)dx$ is a function whose derivative is $f(x)$.

That is to say, $\frac{d}{dx}[\int f(x)dx]=f(x)$.

If $A(x)$ is an indefinite integral of $f(x)$, then so is $A(x)+C$, where C is any constant. Thus, it is customary when evaluating indefinite integrals to add a "constant of integration" C.

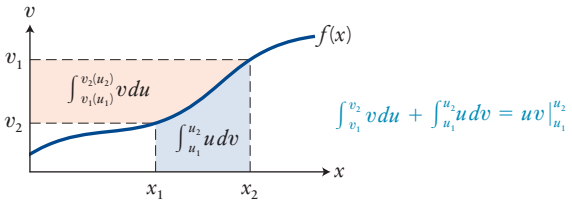
Definite integrals The definite integral of $f(x)$, written as $\int_{x_1}^{x_2}f(x)dx$, represents the sum of the area of contiguous rectangles that each intersect $f(x)$ at some point along one base and that each have another base coincident with the x axis over some part of the range between x_1 and x_2; the indefinite integral evaluates the sum in the limit of arbitrarily small rectangle bases. In other words, the indefinite integral gives the net area that lies under $f(x)$ but above the x axis between the boundaries x_1 and x_2.



If $A(x)$ is any indefinite integral of $f(x)$, then the definite integral is given by $\int_{x_1}^{x_2}f(x)dx=A(x_2)-A(x_1)\equiv A(x)|_{x_1}^{x_2}$. The constant of integration C does not affect the value of definite integrals and thus can be ignored (i.e., set to zero) during evaluation.

Integration by parts $\int_a^b udv$ is the area under the curve of $u(v)$. If $\int_a^b udv$ is difficult to evaluate directly, it is sometimes easier to express the area under the curve as the area within part of a rectangle minus the area under the curve of $v(u)$. In other words:

$$\int_a^b udv=uv|_a^b-\int_a^b vdu.$$



By choosing u and dv appropriately (both can be functions of x), this approach, called "integration by parts," can transform difficult integrals into easier ones.

Table of integrals In the following expressions, a and b are constants. An arbitrary constant of integration C can be added to the right-hand side.

$$\int x^ndx=\frac{1}{n+1}x^{n+1}\ (\text{for } n\neq -1)$$

$$\int x^{-1}dx=\ln|x|$$

$$\int\frac{1}{a^2+x^2}dx=\frac{1}{a}\tan^{-1}\frac{x}{a}$$

$$\int\frac{1}{(a^2+x^2)^2}dx=\frac{1}{2a^3}\tan^{-1}\frac{x}{a}+\frac{x}{2a^2(x^2+a^2)}$$

$$\int\frac{1}{\sqrt{\pm a^2+x^2}}dx=\ln|x+\sqrt{\pm a^2+x^2}|$$

$$\int\frac{1}{\sqrt{a^2-x^2}}dx=\sin^{-1}\frac{x}{|a|}=\tan^{-1}\frac{x}{\sqrt{a^2-x^2}}$$

$$\int\frac{x}{\sqrt{\pm a^2-x^2}}dx=-\sqrt{\pm a^2-x^2}$$

$$\int\frac{x}{\sqrt{\pm a^2+x^2}}dx=\sqrt{\pm a^2+x^2}$$

$$\int\frac{1}{(\pm a^2+x^2)^{3/2}}dx=\frac{\pm x}{a^2\sqrt{\pm a^2+x^2}}$$

$$\int\frac{x}{(a^2+x^2)^{3/2}}dx=-\frac{1}{\sqrt{a^2+x^2}}$$

$$\int\frac{1}{a+bx}dx=\frac{1}{b}\ln(a+bx)$$

$$\int\frac{1}{(a+bx)^2}dx=-\frac{1}{b(a+bx)}$$

$$\int\sin(ax)dx=-\frac{1}{a}\cos(ax)$$

$$\int \cos(ax)dx = \frac{1}{a}\sin(ax)$$

$$\int \tan(ax)dx = -\frac{1}{a}\ln(\cos ax)$$

$$\int \sin^2(ax)dx = \frac{x}{2} - \frac{\sin 2ax}{4a}$$

$$\int \cos^2(ax)dx = \frac{x}{2} + \frac{\sin 2ax}{4a}$$

$$\int x\sin(ax)dx = \frac{1}{a^2}\sin ax - \frac{1}{a}x\cos ax$$

$$\int x\cos(ax)dx = \frac{1}{a^2}\cos ax + \frac{1}{a}x\sin ax$$

$$\int e^{ax}dx = \frac{1}{a}e^{ax}$$

$$\int xe^{ax}dx = \frac{e^{ax}}{a^2}(ax-1)$$

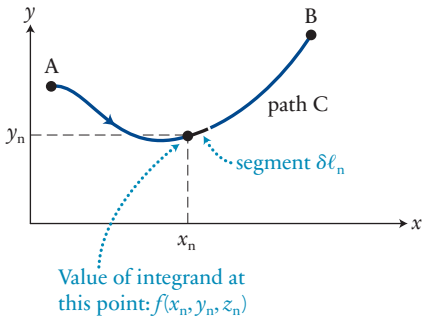
$$\int x^2e^{ax}dx = \frac{x^2e^{ax}}{a} - \frac{2}{a}\left[\frac{e^{ax}}{a^2}(ax-1)\right]$$

$$\int \ln ax\,dx = x\ln(ax) - x$$

$$\int_0^\infty x^n e^{-ax}dx = \frac{n!}{a^{n+1}}$$

$$\int_0^\infty e^{-ax^2}dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}$$

Line integrals A *line integral* is an integral of a function that needs to be evaluated over a path (that is, a curve connecting two points in space). Consider, for example, the two-dimensional path C from point A to point B in the figure below. (The procedure described below is equally applicable in three dimensions.)



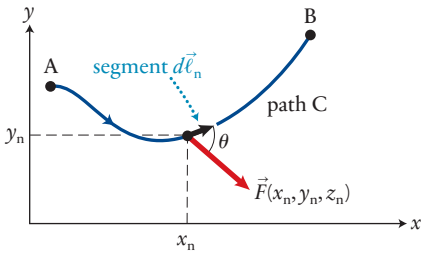
The path from A to B is *directed:* At any point the direction along the path away from A and toward B is forward (positive). Suppose we have a function $f(x, y, z)$ defined everywhere along the path. The function can be either a scalar or a vector; we will first discuss line integrals of scalar functions. We divide the path between A and B into small segments of length $\delta\ell_n$, each segment small enough that we can consider it essentially straight and small enough that the value of the function $f(x, y, z)$ can be considered constant over that segment. We then calculate the product $f(x_n, y_n, z_n)\delta\ell_n$ for each segment. The line integral of the function $f(x, y, z)$ along path C is then given by the sum of all those products along the path in the limit of infinitesimally small segments:

$$\int_C f(x, y, z)d\ell = \lim_{\delta\ell\to\infty}\sum_n f(x_n, y_n, z_n)\delta\ell_n.$$

To evaluate the integral on the right, we need to know the path C. Usually the path is specified in terms of the length parameter ℓ: $x = x(\ell)$, $y = y(\ell)$, $z = z(\ell)$. The line integral can then be written as an ordinary definite integral:

$$\int_C f(x, y, z)d\ell = \int_A^B f[x(\ell), y(\ell), z(\ell)]d\ell.$$

Next we consider the line integral of a vector function. We consider the same path C from A to B, but now we consider a vector function $\vec{F}(x, y, z)$. Instead of taking infinitesimally small scalar segments $d\ell_n$ along the path, we take small vector segments $d\vec{\ell}_n$ along the path of length $d\ell_n$ and whose direction is tangent to the path in the direction of the path from A to B:



At each point we calculate the scalar product $\vec{F}(x_n, y_n, z_n)\cdot d\vec{\ell}_n$ and then sum these products over path C to obtain the line integral.

$$\int_C \vec{F}(x, y, z)\cdot d\vec{\ell}.$$

By writing out the scalar product, $\vec{F}(x, y, z)\cdot d\vec{\ell} = F(x, y, z)\cdot\cos\theta\, d\ell$, we can reduce the line integral of a vector function to that of a scalar function:

$$\int_C \vec{F}(x, y, z)\cdot d\vec{\ell} = \int_C F(x, y, z)\cos\theta\, d\ell.$$

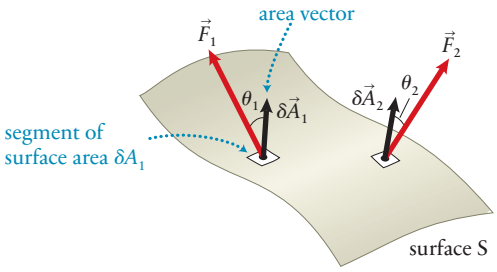
In other words, we need to compute the line integral of the component of the vector $\vec{F}(x, y, z)$ along the tangent to the path.

If the path is closed—that is, the path returns to the starting point—we indicate that by putting a circle through the integration sign:

$$\oint_C \vec{F}(x, y, z)\cdot d\vec{\ell}.$$

Surface integrals A *surface integral* is an integral of a function that needs to be evaluated over a surface. As with line integrals, the integrand of a surface integral can be a scalar or a vector function. We will only discuss the more general case of a vector function here.

The surface over which the integration is to be taken can be either *closed* or *open*. A closed surface, such as the surface of a sphere, divides space into two parts—an inside and an outside—and to get from one part to the other one has to go through the surface. An open surface does not have this property: For the surface S shown in the figure below, for example, one can go from one side of the surface to the other without passing through it.



Consider a vector function $\vec{F}(x, y, z)$. To calculate the surface integral of this function over surface S, we begin by dividing the surface into small segments of surface area δA_n, each segment being small enough that we can consider it to be essentially flat and small enough so that the function $\vec{F}(x, y, z)$ can be considered constant over the segment. We then define an *area vector* $\delta\vec{A}_n$ whose magnitude is equal to the surface area δA_n of the segment and whose direction is normal to that segment. For each segment we then calculate the scalar product of the area vector and the value $\vec{F}_n$ of the vector function at that location: $\vec{F}_n(x_n, y_n, z_n) \cdot \delta\vec{A}_n$. The surface integral of the vector function over the surface S is then given by the sum of all those products for all the segments that make up the surface:

$$\int_S \vec{F}(x, y, z) \cdot d\vec{A} = \lim_{\delta\vec{A}_n \to \infty} \sum_n \vec{F}_n(x_n, y_n, z_n) \cdot \delta\vec{A}_n.$$

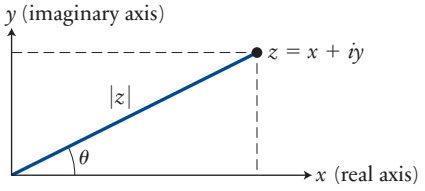
If the surface is closed, we indicate that by putting a circle through the integration sign:

$$\oint_S \vec{F}(x, y, z) \cdot d\vec{A}$$

6 Complex numbers

A complex number $z = x + iy$ is defined in terms of its real part x and its imaginary part y. Both x and y are real numbers. i is Euler's constant, defined by the property $i^2 = -1$.

Each complex number z has a "complex conjugate" z^* which has the same real part but an imaginary part with opposite sign: $z = x + iy \Leftrightarrow z^* = x - iy$.



The real and imaginary parts can be expressed in terms of the complex number and its conjugate:

$$x = \tfrac{1}{2}(z + z^*)$$

$$y = \tfrac{1}{2}i(z - z^*)$$

A complex number is like a two-dimensional vector in a plane with a real axis and an imaginary axis. Thus, z can be described by a magnitude or length $|z|$ and an angle θ formed with the real axis (called the "phase angle"):

$$z = |z|(\cos\theta + i\sin\theta), \text{ where } |z| = \sqrt{zz^*} \text{ and}$$

$$\theta = \tan^{-1}\frac{y}{x} = \tan^{-1}\frac{i(z - z^*)}{(z + z^*)}.$$

Euler's formula says that $e^{i\theta} = \cos\theta + i\sin\theta$, allowing complex numbers to be written in the form $z = |z|e^{i\theta}$. This is a convenient form for expressing complex numbers. For example, it is easy to raise a complex number z to a power n: $z^n = |z|^n e^{in\theta}$.

7 Useful approximations

Binomial expansion

$$(1 + x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \cdots$$

If $x \ll 1$, then $(1 + x)^n \approx 1 + nx$

Trigonometric expansions

$$\sin\alpha = \alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \frac{\alpha^7}{7!} + \cdots \quad (\alpha \text{ in rad})$$

$$\cos\alpha = 1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} - \frac{\alpha^6}{6!} + \cdots \quad (\alpha \text{ in rad})$$

$$\tan\alpha = \alpha + \frac{1}{3}\alpha^3 + \frac{2}{15}\alpha^5 + \frac{17}{315}\alpha^7 + \cdots \quad (\alpha \text{ in rad})$$

If $\alpha \ll 1$ rad, then $\sin\alpha \approx \alpha$, $\cos\alpha \approx 1$, and $\tan\alpha \approx \alpha$.

Other useful expansions

$$\frac{1}{1 - x} = 1 + x + x^2 + x^3 + \cdots \text{ for } -1 < x < 1$$

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \cdots$$

$$\ln(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots \text{ for } -1 < x < 1$$

$$\ln\left(\frac{1 + x}{1 - x}\right) = 2x + \frac{2}{3}x^3 + \frac{2}{5}x^5 - \frac{2}{7}x^7 + \cdots \text{ for } -1 < x < 1$$

Appendix C

SI Units, Useful Data, and Unit Conversion Factors

The seven base SI units

Unit	Abbreviation	Physical quantity
meter	m	length
kilogram	kg	mass
second	s	time
ampere	A	electric current
kelvin	K	thermodynamic temperature
mole	mol	amount of substance
candela	cd	luminous intensity

Some derived SI units

Unit	Abbreviation	Physical quantity	In terms of base units
newton	N	force	$kg \cdot m/s^2$
joule	J	energy	$kg \cdot m^2/s^2$
watt	W	power	$kg \cdot m^2/s^3$
pascal	Pa	pressure	$kg/m \cdot s^2$
hertz	Hz	frequency	s^{-1}
coulomb	C	electric charge	$A \cdot s$
volt	V	electric potential	$kg \cdot m^2/(A \cdot s^3)$
ohm	Ω	electric resistance	$kg \cdot m^2/(A^2 \cdot s^3)$
farad	F	capacitance	$A^2 \cdot s^4/(kg \cdot m^2)$
tesla	T	magnetic field	$kg/(A \cdot s^2)$
weber	Wb	magnetic flux	$kg \cdot m^2/(A \cdot s^2)$
henry	H	inductance	$kg \cdot m^2/(A^2 \cdot s^2)$

SI Prefixes

10^n	Prefix	Abbreviation	10^n	Prefix	Abbreviation
10^0	—	—			
10^3	kilo-	k	10^{-3}	milli-	m
10^6	mega-	M	10^{-6}	micro-	μ
10^9	giga-	G	10^{-9}	nano-	n
10^{12}	tera-	T	10^{-12}	pico-	p
10^{15}	peta-	P	10^{-15}	femto-	f
10^{18}	exa-	E	10^{-18}	atto-	a
10^{21}	zetta-	Z	10^{-21}	zepto-	z
10^{24}	yotta-	Y	10^{-24}	yocto-	y

Values of fundamental constants

Quantity	Symbol	Value
Speed of light in vacuum	c_0	3.00×10^8 m/s
Gravitational constant	G	6.6738×10^{-11} N · m^2/kg^2
Avogadro's number	N_A	6.0221413×10^{23} mol^{-1}
Boltzmann's constant	k_B	1.381×10^{-23} J/K
Charge on electron	e	1.60×10^{-19} C
Permittivity constant	ϵ_0	$8.85418782 \times 10^{-12}$ C^2/(N · m^2)
Permeability constant	μ_0	$4\pi \times 10^{-7}$ T · m/A
Planck's constant	h	6.626×10^{-34} J · s
Electron mass	m_e	9.11×10^{-31} kg
Proton mass	m_p	1.6726×10^{-27} kg
Neutron mass	m_n	1.6749×10^{-27} kg
Atomic mass unit	amu	1.6605×10^{-27} kg

Other useful numbers

Number or quantity	Value
π	3.1415927
e	2.7182818
1 radian	57.2957795°
Absolute zero ($T = 0$)	−273.15 °C
Average acceleration g due to gravity near Earth's surface	9.8 m/s^2
Speed of sound in air at 20 °C	343 m/s
Density of dry air at atmospheric pressure and 20 °C	1.29 kg/m^3
Earth's mass	5.97×10^{24} kg
Earth's radius (mean)	6.38×10^6 m
Earth–Moon distance (mean)	3.84×10^8 m

Unit conversion factors

Length

1 in. = 2.54 cm (defined)

1 cm = 0.3937 in.

1 ft = 30.48 cm

1 m = 39.37 in. = 3.281 ft

1 mi = 5280 ft = 1.609 km

1 km = 0.6214 mi

1 nautical mile (U.S.) = 1.151 mi = 6076 ft = 1.852 km

1 fermi = 1 femtometer (fm) = 10^{-15} m

1 angstrom (Å) = 10^{-10} m = 0.1 nm

1 light-year (ly) = 9.461×10^{15} m

1 parsec = 3.26 ly = 3.09×10^{16} m

Volume

1 liter (L) = 1000 mL = 1000 cm^3 = 1.0×10^{-3} m^3
= 1.057 qt (U.S.) = 61.02 $in.^3$

1 gal (U.S.) = 4 qt (U.S.) = 231 $in.^3$ = 3.785 L = 0.8327 gal (British)

1 quart (U.S.) = 2 pints (U.S.) = 946 mL

1 pint (British) = 1.20 pints (U.S.) = 568 mL

1 m^3 = 35.31 ft^3

Speed

1 mi/h = 1.4667 ft/s = 1.6093 km/h = 0.4470 m/s

1 km/h = 0.2778 m/s = 0.6214 mi/h

1 ft/s = 0.3048 m/s = 0.6818 mi/h = 1.0973 km/h

1 m/s = 3.281 ft/s = 3.600 km/h = 2.237 mi/h

1 knot = 1.151 mi/h = 0.5144 m/s

Angle

1 radian (rad) = 57.30° = 57°18'

1° = 0.01745 rad

1 rev/min (rpm) = 0.1047 rad/s

Time

1 day = 8.640×10^4 s

1 year = 365.242 days = 3.156×10^7 s

Mass

1 atomic mass unit (u) = 1.6605×10^{-27} kg

1 kg = 0.06852 slug

1 metric ton = 1000 kg

1 long ton = 2240 lb = 1016 kg

1 short ton = 2000 lb = 909.1 kg

1 kg has a weight of 2.20 lb where g = 9.80 m/s^2

Force

1 lb = 4.44822 N

1 N = 10^5 dyne = 0.2248 lb

Energy and work

1 J = 10^7 ergs = 0.7376 ft · lb

1 ft · lb = 1.356 J = 1.29×10^{-3} Btu = 3.24×10^{-4} kcal

1 kcal = 4.19×10^3 J = 3.97 Btu

1 eV = 1.6022×10^{-19} J

1 kWh = 3.600×10^6 J = 860 kcal

1 Btu = 1.056×10^3 J

Power

1 W = 1 J/s = 0.7376 ft · lb = 3.41 Btu/h

1 hp = 550 ft · lb = 746 W

1 kWh/day = 41.667 W

Pressure

1 atm = 1.01325 bar = 1.01325×10^5 N/m^2 = 14.7 $lb/in.^2$ = 760 torr

1 $lb/in.^2$ = 6.895×10^3 N/m^2

1 Pa = 1 N/m^2 = 1.450×10^{-4} $lb/in.^2$

Periodic Table of the Elements

Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1 H 1.008																	2 He 4.003
2	3 Li 6.941	4 Be 9.012											5 B 10.811	6 C 12.011	7 N 14.007	8 O 15.999	9 F 18.998	10 Ne 20.180
3	11 Na 22.990	12 Mg 24.305											13 Al 26.982	14 Si 28.086	15 P 30.974	16 S 32.065	17 Cl 35.453	18 Ar 39.948
4	19 K 39.098	20 Ca 40.078	21 Sc 44.956	22 Ti 47.867	23 V 50.942	24 Cr 51.996	25 Mn 54.938	26 Fe 55.845	27 Co 58.933	28 Ni 58.693	29 Cu 63.546	30 Zn 65.409	31 Ga 69.723	32 Ge 72.64	33 As 74.922	34 Se 78.96	35 Br 79.904	36 Kr 83.798
5	37 Rb 85.468	38 Sr 87.62	39 Y 88.906	40 Zr 91.224	41 Nb 92.906	42 Mo 95.94	43 Tc (98)	44 Ru 101.07	45 Rh 102.906	46 Pd 106.42	47 Ag 107.868	48 Cd 112.411	49 In 114.818	50 Sn 118.710	51 Sb 121.760	52 Te 127.60	53 I 126.904	54 Xe 131.293
6	55 Cs 132.905	56 Ba 137.327	71 Lu 174.967	72 Hf 178.49	73 Ta 180.948	74 W 183.84	75 Re 186.207	76 Os 190.23	77 Ir 192.217	78 Pt 195.078	79 Au 196.967	80 Hg 200.59	81 Tl 204.383	82 Pb 207.2	83 Bi 208.980	84 Po (209)	85 At (210)	86 Rn (222)
7	87 Fr (223)	88 Ra (226)	103 Lr (262)	104 Rf (261)	105 Db (262)	106 Sg (266)	107 Bh (264)	108 Hs (269)	109 Mt (268)	110 Ds (271)	111 Rg (272)	112 Cn (285)	113 Nh (284)	114 Fl (289)	115 Mc (289)	116 Lv (292)	117 Ts (294)	118 Og (294)

Period

Number of protons → 29
Symbol for element → Cu
63.546 ← Average atomic mass in g/mol. For elements having no stable isotope, value in parentheses is approximate atomic mass of longest-lived isotope.

Lanthanoids	57 La 138.905	58 Ce 140.116	59 Pr 140.908	60 Nd 144.24	61 Pm (145)	62 Sm 150.36	63 Eu 151.964	64 Gd 157.25	65 Tb 158.925	66 Dy 162.500	67 Ho 164.930	68 Er 167.259	69 Tm 168.934	70 Yb 173.04
Actinoids	89 Ac (227)	90 Th (232)	91 Pa (231)	92 U (238)	93 Np (237)	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)

Appendix D

Center of Mass of Extended Objects

We can apply the concept of center of mass to extended objects. Consider, for example, the object of inertia m in **Figure D.1**. If you imagine breaking down the object into many small segments of equal inertia δm, you can use Eq. 6.24 to compute the position of the center of mass:

$$x_{\text{cm}} = \frac{\delta m_1 x_1 + \delta m_2 x_2 + \cdots}{\delta m_1 + \delta m_2 + \cdots}, \tag{D.1}$$

where x_n is the position of segment δm_n. Because the sum of the inertias of all segments is equal to the inertia m of the extended object, $\delta m_1 + \delta m_2 + \cdots = m$, we can write Eq. D.1 as

$$x_{\text{cm}} = \frac{1}{m}(\delta m_1 x_1 + \delta m_2 x_2 + \cdots) = \frac{1}{m}\sum_n (\delta m_n x_n) \tag{D.2}$$

To evaluate this sum for the extended object, we take the limit of this expression as $\delta m \to 0$. In this limit, the sum becomes an integral:

$$x_{\text{cm}} = \frac{1}{m}\lim_{\delta m \to 0}\sum_n (\delta m_n x_n) \equiv \frac{1}{m}\int_{\text{object}} x\, dm. \tag{D.3}$$

To evaluate the integral we need to know how the inertia is distributed over the object. Let the *inertia per unit length* of the object be $\lambda \equiv dm/dx$. In general λ is a function of position—that is, the inertia per unit length need not be the same at different locations along the object—and so $\lambda = \lambda(x)$ and $dm = \lambda(x)dx$. Substituting this expression for dm into Eq. D.3, we obtain

$$x_{\text{cm}} = \frac{1}{m}\int_{\text{object}} x\lambda(x)dx. \tag{D.4}$$

In this expression, the limits of integration should be taken to be the positions of the left and right ends of the object (x_{L} and x_{R}, respectively).

For example, let the object in Figure D.1 have a uniformly distributed inertia so that $\lambda(x)$ has the same value (the inertia of the extended object divided by the length of the object) everywhere:

$$\lambda(x) = \frac{m}{x_{\text{R}} - x_{\text{L}}}. \tag{D.5}$$

Because $\lambda(x)$ does not depend on x, we can pull it out of the integral in Eq. D.4 and so, substituting Eq. D.5 into Eq. D.4, we get

$$x_{\text{cm}} = \frac{\lambda(x)}{m}\int_{x_{\text{L}}}^{x_{\text{R}}} x\, dx = \frac{1}{x_{\text{R}} - x_{\text{L}}}\int_{x_{\text{L}}}^{x_{\text{R}}} x\, dx = \frac{1}{x_{\text{R}} - x_{\text{L}}}\left[\tfrac{1}{2}x^2\right]_{x_{\text{L}}}^{x_{\text{R}}}$$

$$= \frac{x_{\text{R}}^2 - x_{\text{L}}^2}{2(x_{\text{R}} - x_{\text{L}})} = \frac{x_{\text{R}} + x_{\text{L}}}{2}. \tag{D.6}$$

In other words, the center of mass of the object is halfway between the ends of the object (at the center of the object), as we expect.

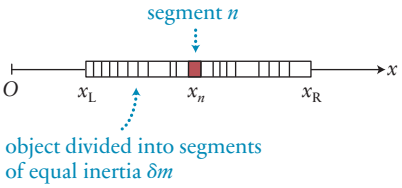


Figure D.1

Appendix E

Derivation of the Lorentz Transformation Equations

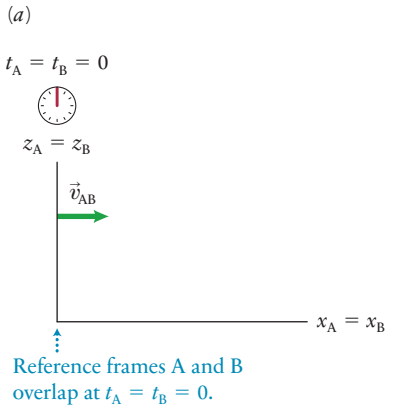


Figure E.1 Determining the space and time coordinates of an event from two reference frames that overlap at $t_A = t_B = 0$.

In this appendix we derive the Lorentz transformation equations (Eqs. 14.29–14.32), which relate position and time in two reference frames A and B. The origins of the reference frames coincide at $t_A = t_B = 0$, and the frames are moving at constant velocity $\vec{v}_{AB} = v_{ABx}\hat{i}$ in the x direction relative to each other (**Figure E.1*a***).

Consider an event e that occurs at instant t_{Ae} and position x_{Ae} in reference frame A (Figure E.1*b*).* Our goal is to express t_{Be} and x_{Be}, the instant and the position at which the event occurs in reference frame B, in terms of t_{Ae} and x_{Ae}:

$$t_{Be} = f(x_{Ae}, t_{Ae}) \tag{E.1}$$

$$x_{Be} = g(x_{Ae}, t_{Ae}). \tag{E.2}$$

To obtain these relationships, we consider the space-time interval (Eq. 14.18) between the event in Figure E.1*b* and the event in Figure E.1*a* (this event being the overlapping of the two origins). The spatial separation of these two events is $\Delta x_{Ae} = x_{Ae} - 0 = x_{Ae}$ in reference frame A and $\Delta x_{Be} = x_{Be} - 0 = x_{Be}$ in reference frame B. The temporal separations are $\Delta t_{Ae} = t_{Ae} - 0 = t_{Ae}$ and $\Delta t_{Be} = t_{Be} - 0 = t_{Be}$. Because the space-time interval is an invariant, we have

$$(c_0 t_{Be})^2 - (x_{Be})^2 = (c_0 t_{Ae})^2 - (x_{Ae})^2. \tag{E.3}$$

To satisfy Eq. E.3, the relationships expressed in Eqs. E.1 and E.2 must be linear:

$$t_{Be} = Dt_{Ae} + Ex_{Ae} \tag{E.4}$$

$$x_{Be} = Ft_{Ae} + Gx_{Ae}, \tag{E.5}$$

where D, E, F, and G are constants. To obtain relationships between these constants, consider the position of the origin of reference frame B at the instant illustrated in Figure E.1*b*. The origin of reference frame B moves at constant velocity $\vec{v}_{AB}$ relative to reference frame A, and so at t_{Ae} the position of B's origin in reference frame A is

$$x_{AB} = v_{ABx}t_{Ae}. \tag{E.6}$$

The position of B's origin in reference frame B is always zero, $x_{BB} = 0$.† According to Eq. E.5 we then have $x_{BB} = 0 = Ft_{Ae} + Gx_{AB}$, or substituting for x_{AB} from Eq. E.6, $0 = Ft_{Ae} + G(v_{ABx}t_{Ae})$. Therefore

$$F = -v_{ABx}G. \tag{E.7}$$

Conversely, the origin of reference frame A is located at $x_{BA} = -v_{ABx}t_{Be}$ in reference frame B, and so, with $x_{AA} = 0$, Eqs. E.4 and E.5 become

$$t_{Be} = Dt_{Ae} + 0 \tag{E.8}$$

$$-v_{ABx}t_{Be} = Ft_{Ae} + 0. \tag{E.9}$$

Solving Eqs. E.8 and E.9 for F gives $F = -v_{ABx}D$, and so, from Eq. E.7,

$$D = G. \tag{E.10}$$

*Remember our subscript format: The capital letter refers to the reference frame; the lowercase e is for "event." Thus the vector $\vec{r}_{Ae}$ represents observer **A**'s measurement of the position at which the event occurs.

†Do not be stymied by the unfamiliar subscript BB. It's the same format we've been using all along: The first B tells you the reference frame in which the measurement is made, and the second B tells you what's being measured—in this case, the position of the origin of reference frame B.

Equations E.4 and E.5 then become

$$t_{\mathrm{Be}} = Gt_{\mathrm{Ae}} + Ex_{\mathrm{Ae}} \tag{E.11}$$

$$x_{\mathrm{Be}} = -v_{\mathrm{ABx}}Gt_{\mathrm{Ae}} + Gx_{\mathrm{Ae}}. \tag{E.12}$$

Substituting Eqs. E.11 and E.12 into the left side of Eq. E.3 and using $v_{\mathrm{ABx}}^2 = v^2$ yields

$$c_0^2G^2\left(1 - \frac{v^2}{c_0^2}\right)t_{\mathrm{Ae}}^2 + (c_0^2EG + G^2v_{\mathrm{ABx}})2t_{\mathrm{Ae}}x_{\mathrm{Ae}} - (G^2 - c_0^2E^2)x_{\mathrm{Ae}}^2$$
$$= (c_0t_{\mathrm{Ae}})^2 - (x_{\mathrm{Ae}})^2. \tag{E.13}$$

and then gather all the $(t_{\mathrm{Ae}})^2$ terms, all the $x_{\mathrm{Ae}}\,t_{\mathrm{Ae}}$ terms, and all the $(t_{\mathrm{Ae}})^2$ terms:

$$\left[c_0^2G^2\left(1 - \frac{v^2}{c_0^2}\right) - c_0^2\right]t_{\mathrm{Ae}}^2 + 2(c_0^2EG + G^2v_{\mathrm{ABx}})\,x_{\mathrm{Ae}}t_{\mathrm{Ae}}$$
$$-(G^2 - c_0^2E^2 - 1)x_{\mathrm{Ae}}^2 = 0. \tag{E.14}$$

Because Eq. E.14 must hold for any value of t_{Ae} and x_{Ae}, the coefficient of each term must be zero. The coefficient of the t_{Ae}^2 term yields

$$c_0^2G^2\left(1 - \frac{v^2}{c_0^2}\right) = c_0^2. \tag{E.15}$$

Solving Eq. E.15 for G yields

$$G = \left(1 - \frac{v^2}{c_0^2}\right)^{-1/2} \equiv \gamma, \tag{E.16}$$

and so from Eq. E.7, $F = -v_{\mathrm{ABx}}G$, we see that $F = -\gamma v_{\mathrm{ABx}}$.

The coefficient of the $x_{\mathrm{Ae}}t_{\mathrm{Ae}}$ term in Eq. E.14 must also be zero. Substituting γ for G in that coefficient gives us

$$c_0^2E\gamma + \gamma^2v_{\mathrm{ABx}} = 0 \tag{E.17}$$

$$E = -\gamma v_{\mathrm{ABx}}/c_0^2. \tag{E.18}$$

Now we are ready to substitute for *D*, *E*, *F*, and *G* in Eqs. E.4 and E.5—*D* from Eq. E.10, *E* from Eq. E.18, *F* from Eq. E.7, and *G* from Eq. E.16:

$$t_{\mathrm{Be}} = \gamma\left(t_{\mathrm{Ae}} - \frac{1}{c_0^2}v_{\mathrm{ABx}}x_{\mathrm{Ae}}\right) \tag{E.19}$$

$$x_{\mathrm{Be}} = \gamma(x_{\mathrm{Ae}} - v_{\mathrm{ABx}}t_{\mathrm{Ae}}). \tag{E.20}$$

Because there is no length contraction in the directions perpendicular to the relative velocity $\vec{v}_{\mathrm{AB}}$ of the two reference frames, we have

$$y_{\mathrm{Be}} = y_{\mathrm{Ae}} \tag{E.21}$$

$$z_{\mathrm{Be}} = z_{\mathrm{Ae}}. \tag{E.22}$$

Equations E.19–E.22 are the Lorentz transformation equations we wanted to derive.

Velocities in two reference frames

We can use Eqs. E.19 and E.20 to derive Eq. 14.33, the relationship between the *x* components of the velocities of an object o measured in two reference frames A and B. Let the *x* component be $v_{\mathrm{Aox}} = dx_{\mathrm{Ao}}/dt_{\mathrm{A}}$ in reference frame A and $v_{\mathrm{Box}} = dx_{\mathrm{Bo}}/dt_{\mathrm{B}}$ in reference frame B. We begin by writing Eqs. E.19 and E.20 in differential form:

$$dt_{\mathrm{B}} = \gamma\left(dt_{\mathrm{A}} - \frac{v_{\mathrm{ABx}}}{c_0^2}dx_{\mathrm{Ao}}\right) \tag{E.23}$$

$$dx_{\mathrm{Bo}} = \gamma(dx_{\mathrm{Ao}} - v_{\mathrm{ABx}}\,dt_{\mathrm{A}}), \tag{E.24}$$

$$dy_{\mathrm{Bo}} = dy_{\mathrm{Ao}}, \tag{E.25}$$

$$dz_{\mathrm{Bo}} = dz_{\mathrm{Ao}}, \tag{E.26}$$

where we have replaced the subscript e by o because we are considering an object rather than an event. Dividing Eq. E.24 by Eq. E.23 yields an expression for the x component of the velocity in reference frame B:

$$v_{\mathrm{B}ox} = \frac{dx_{\mathrm{B}o}}{dt_{\mathrm{B}}} = \frac{\gamma(dx_{\mathrm{A}o} - v_{\mathrm{AB}x}\,dt_{\mathrm{A}})}{\gamma\left(dt_{\mathrm{A}} - \dfrac{v_{\mathrm{AB}x}}{c_0^2}dx_{\mathrm{A}o}\right)} = \frac{(dx_{\mathrm{A}o} - v_{\mathrm{AB}x}\,dt_{\mathrm{A}})}{\left(dt_{\mathrm{A}} - \dfrac{v_{\mathrm{AB}x}}{c_0^2}dx_{\mathrm{A}o}\right)}. \tag{E.27}$$

Finally, we divide the numerator and denominator of the rightmost term by dt_{A} to obtain Eq. 14.33:

$$v_{\mathrm{B}ox} = \frac{\left(\dfrac{dx_{\mathrm{A}o}}{dt_{\mathrm{A}}} - v_{\mathrm{AB}x}\right)}{\left(1 - \dfrac{v_{\mathrm{AB}x}}{c_0^2}\dfrac{dx_{\mathrm{A}o}}{dt_{\mathrm{A}}}\right)} = \frac{(v_{\mathrm{A}ox} - v_{\mathrm{AB}x})}{\left(1 - \dfrac{v_{\mathrm{AB}x}}{c_0^2}v_{\mathrm{A}ox}\right)}. \tag{E.28}$$

To find the y component of the velocity in reference frame B, we divide Eq. E.25 by Eq. E.23:

$$v_{\mathrm{B}oy} = \frac{dy_{\mathrm{B}o}}{dt_{\mathrm{B}}} = \frac{dy_{\mathrm{A}o}}{\gamma\left(dt_{\mathrm{A}} - \dfrac{v_{\mathrm{AB}x}}{c_0^2}dx_{\mathrm{A}o}\right)}. \tag{E.29}$$

Dividing numerator and denominator by dt_{A} then yields

$$v_{\mathrm{B}oy} = \frac{v_{\mathrm{A}oy}}{\gamma\left(1 - \dfrac{v_{\mathrm{AB}x}v_{\mathrm{A}ox}}{c_0^2}\right)}. \tag{E.30}$$

For the z component of the velocity we obtain a similar equation with y replaced by z:

$$v_{\mathrm{B}oz} = \frac{v_{\mathrm{A}oz}}{\gamma\left(1 - \dfrac{v_{\mathrm{AB}x}v_{\mathrm{A}ox}}{c_0^2}\right)}. \tag{E.31}$$

$$\Delta p_{sx} = +0.38\ \text{kg} \cdot \text{m/s} - 0$$

Solutions to Checkpoints

Chapter 1

1.1 A nickel and a quarter. If you had to think about this for more than a few seconds, you probably unknowingly assumed that *neither* coin could be a nickel. This is an example of a hidden assumption.

1.2 That the clock works with the new batteries does not necessarily mean the old batteries are dead. They could have been inserted incorrectly, or there could have been a bad contact that got fixed when the new batteries were inserted. That the clock does not work with the new batteries does not necessarily mean the clock is broken. Both the new and the old batteries could be inserted incorrectly, or both the new and the old batteries could be dead.

1.3 Two valuable goals are to become a better problem solver and to learn to apply the scientific method to any type of reasoning. Either skill is useful in any type of career. If you wrote "pass physics requirement" or "get into medical school," think again. With those goals, it is going to be very difficult for you to derive much satisfaction or benefit from this course.

Here are some interesting responses I have collected over the years: to become a responsible citizen of the world, to understand modern technology, to bridge the gap between the sciences and the arts/humanities, to sharpen my mind.

1.4 Likely answers are that something about the case or cap has changed. The case opening may have expanded due to the heat, or the cap may have acquired a dent that prevents it from fitting over the opening. If you answered something along these lines, you assumed symmetry with respect to translation in both time and space: If the cap fits at home, it ought to fit elsewhere now and anytime in the future.

1.5 The distance from the North Pole to the equator is one-fourth of Earth's circumference c, which means $c = 4 \times (10{,}000{,}000\ \text{m}) = 40{,}000{,}000\ \text{m}$. Because $c = 2\pi R_E$, where R_E is Earth's radius, you have $R_E = c/2\pi = (40{,}000{,}000\ \text{m})/2\pi = 6{,}400{,}000\ \text{m}$.

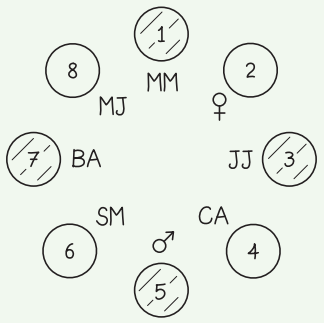
1.6 The diameter of an atom is 10^{-10} m (see Figure 1.9); the diameter of an apple is about 10^{-1} m. Therefore the required magnification factor is $(10^{-1}\ \text{m})/(10^{-10}\ \text{m}) = 10^9$. With this magnification, the apple's diameter would be $10^9(10^{-1}\ \text{m}) = 10^8$ m, which is about 10 times Earth's diameter.

1.7 (*a*) (*i*) Somebody switches the power off; the light bulb burns out. (*ii*) Thunder; tractor trailer passing dorm; stomach growling; dorm starting to collapse. (*iii*) You spent all the money in your account; you forgot to deposit the money your parents gave you last week; the bank made a mistake. (*b*) No. All these causes occur before the event. (*c*) One emotion you almost certainly experience is anxiety. Such events tend to be unsettling—we want to know the causes of things we observe.

1.8 A typical physics class lasts about 1 h, which is 3600 s, or order of magnitude 10^4 s. The number of sequential chemical reactions is thus $(10^4\ \text{s})(1\ \text{reaction}/10^{-13}\ \text{s}) = 10^{17}$ (hundred million billion!) reactions.

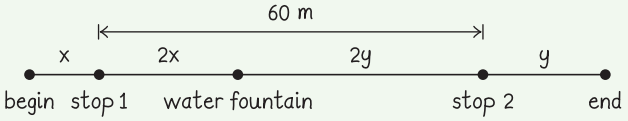
1.9 Because Mike Masters imposes the most constraints, you begin by placing him in seat 1 (Figure S1.9). To satisfy the constraint of having men and women alternate, placing Mike here means the odd-numbered seats are for the men. Only two arrangements avoid putting Cyndi Ahlers next to Mike: Put her in 4 or in 6. If you put Cyndi in 4, Sylvia Masters has to go in 6 to avoid being next to her husband in 1. With Sylvia in 6, John Jones cannot be in 5 or 7, and so you have to put him in 3. Mary Jones then has to go in 8. The remaining empty seats—2, 5, and 7—leave you no choice: Bob Ahlers has to go in 7 because putting him in 5 would have him next to his wife in 4, and your spouse and you get 2 and 5. (Had you put Cyndi Ahlers in 6, you would have obtained the mirror image of the arrangement shown in Figure S1.9.)

Figure S1.9



1.10 90 m. If you call the distance from the west end of the corridor to stop 1 x, the distance from stop 1 to the fountain is $2x$ (Figure S1.10). If you call the distance from stop 2 to the fountain $2y$, the distance from stop 2 to the east end of the corridor is y. Because the distance between the two stops is 60 m, you know that $2x + 2y = 60$ m. From the figure, you see that the total corridor length L is $L = x + 2x + 2y + y = 3x + 3y$. The rest is algebra: $2x + 2y = 60$ m, $x = (30\ \text{m} - y)$, $L = 3(30\ \text{m} - y) + 3y = 90$ m.

Figure S1.10



1.11 (*a*) $\ell = 150$ Mm (megameters) or 0.15 Gm (gigameters). (*b*) $t = 12$ ps (picoseconds). (*c*) 1.2 Mm/s. (*d*) 2.3 Mg.

1.12 The ratio on the left has inches in the numerator and millimeters in the denominator, meaning that the inches don't cancel out. In general, when converting from a given unit to a desired unit, you need to use a ratio that has the desired unit in the numerator and the given unit in the denominator.

1.13 (*a*) In 1 mol you have 6.022×10^{23} atoms, and when you pack this number of atoms into a cube, each side of the cube contains $\sqrt[3]{6.022 \times 10^{23}}$ atoms $= 8.445 \times 10^7$ atoms. The diameter of each atom is 10^{-10} m (see Figure 1.9), and so the length of each side of the cube is $(8.445 \times 10^7\ \text{atoms})(10^{-10}\ \text{m/atom}) = 8.4 \times 10^{-3}$ m, which is about one-third of an inch. (*b*) Because 1 mol of carbon has a mass of 12×10^{-3} kg, Eq. 1.4 gives $V = m/\rho = (12 \times 10^{-3}\ \text{kg})/(2.2 \times 10^3\ \text{kg/m}^3) = 5.5 \times 10^{-6}\ \text{m}^3$. If you assume this value to be a cube, then the cube root of V gives the length of a side, 1.8×10^{-2} m, about twice the length for each side calculated in part *a*.

1.14 (*a*) The circumference is $2\pi R = 2\pi(27.3\ \text{mm}) = 172$ mm. You are allowed to report three significant digits because the radius R has three (the 2 in $2\pi R$ has an infinite number of significant digits, and π has as many as your calculator has digits). (*b*) $a + b + c = 12.3 + 3.241 + 55.74 = 71.3$ (12.3 has the fewest decimal places (one), and so the sum must be rounded to one decimal place). (*c*) $f = (4.00)^2/7 = 2.29$; $g = (3.00)^2/7 = 1.29$; $f + g = (4.00)^2/7 + (3.00)^2/7 = 3.57$. Note that this result is not equal to the sum of the rounded values for f and g: $2.29 + 1.29 = 3.58$. Never use a rounded intermediate answer in a calculation!

1.15 Many answers are possible. One translation strategy would be to reason that water is like milk, and you know from experience that a quart of milk weighs about 2 lb. From Appendix C, you know that 1 qt $= 0.946 \times 10^{-3}\ \text{m}^3 \approx 10^{-3}\ \text{m}^3$ and that 2 lb $= 0.907$ kg ≈ 1 kg. Using Eq. 1.4, you thus obtain for the density of milk (and of water) $(1\ \text{kg})/(10^{-3}\ \text{m}^3) = 10^3\ \text{kg/m}^3$.

One division strategy would be to estimate the mass and the volume of a glass of water and then use Eq. 1.4 to obtain the density. A glass of water is about 2 in. across and 4 in. tall, so $V \approx \pi[(1\ \text{in.})(2.54 \times 10^{-2}\ \text{m/in.})]^2(4\ \text{in.})(2.54 \times 10^{-2}\ \text{m/in.}) = 2 \times 10^{-4}\ \text{m}^3$. Such a glass weighs about 0.25 lb, and so $m \approx (0.25\ \text{lb})(0.45\ \text{kg/lb}) = 0.11$ kg. Substituting V and m into Eq. 1.4 yields $\rho = m/V \approx (0.11\ \text{kg})/(2 \times 10^{-4}\ \text{m}^3) = 5.5 \times 10^2\ \text{kg/m}^3$. The order of magnitude of both estimates is $10^3\ \text{kg/m}^3$ because $5.5 \times 10^2\ \text{kg/m}^3$ rounds to $10 \times 10^2\ \text{kg/m}^3 = 10^3\ \text{kg/m}^3$, which is the same answer we obtained before.

1.16 If we ignore the oceans, which constitute only a small layer at Earth's surface, the planet is mostly rock, which has a mass density higher than that of water—say, five times higher. As discussed in Section 1.9, the mass density of water is $1000\ \text{kg/m}^3$, meaning that rock should have a mass density of about $5 \times 10^3\ \text{kg/m}^3$. From Checkpoint 1.5 (which you worked out, didn't you?), you know that the Earth's radius is 6.4×10^6 m. Taking Earth to be a sphere, you obtain for its volume $V = \frac{4}{3}\pi R_E^3 = 1.1 \times 10^{21}\ \text{m}^3$. Rearranging Eq. 1.4 to $m = \rho V$ gives you $m = (5 \times 10^3\ \text{kg/m}^3)(1.1 \times 10^{21}\ \text{m}^3) = 5 \times 10^{24}$ kg, with order of magnitude 10^{25} kg. (The current best measurement of Earth's mass is 5.9736×10^{24} kg.)

Chapter 2

2.1 (*a*) See Table 2.1. (*b*) See Figure 2.2.

2.2 (*a*) Staying the same distance from the left edge in five sequential frames means I have stopped and am standing in one spot as time passes and the camera continues to roll. (*b*) I am again in the same spot in two frames, but now the frames are not sequential, and the data points in between are not aligned, telling you that in the time that passed between frames 7 and 14, I walked away from the spot and then came back to it. (*c*) Two vertically aligned points would mean that in one frame (which means one instant in time), I was at two different distances from the origin, something that could never happen.

2.3 (*a*) My image is 4.5 mm tall in the film clip. Estimating my height to be 1.8 m (6 ft), you get that 4.5 mm in the photograph corresponds to 1.8 m in the real world. So 1 mm in the photograph corresponds to (1 mm in photo)(1.8 m in real world)/(4.5 mm in photo) = 0.40 m in the real world. (*b*) In frame 1, I am 2.5 mm from the left edge of the frame; in frame 10, 12.0 mm. So the distance traveled between frames 1 and 10 is (12.0 mm in photo) − (2.5 mm in photo) = 9.5 mm in photo. This corresponds to an actual distance of (9.5 mm in photo) · (0.40 m in real world)/(1 mm in photo) = 3.8 m.

2.4 (*a*) +3.4 m in Figure 2.3; +2.4 m in Figure 2.4. (*b*) +3.4 m − (+1.0 m) = +2.4 m; +2.4 m − 0 = +2.4 m.

2.5 (*a*) Because your final position is the same as your initial position, the x component of your displacement is zero. (*b*) Your

distance traveled is 4 m. (*c*) No. In general the *distance traveled* and the *x component of the displacement* are two different quantities.

2.6 (*a*) A vertical line drawn from the first $x = +4.0$ m point down to the horizontal axis tells you that I was at that position at 1.7 s. The vertical line for the $x = +1$ m position is already drawn: It is the vertical axis. Therefore the length of time I took to travel this distance was $(1.7\text{ s}) - (0) = 1.7$ s. (*b*) Vertical lines through the curve at $x = +2.0$ m and $x = +3.0$ m intersect the time axis at 0.6 s and 1.2 s, and so I traveled this distance in $(1.2\text{ s}) - (0.6\text{ s}) = 0.6$ s. (*c*) At 0.8 s. (*d*) I was at $x = +2.5$ m for *zero* time (not even a split second, an infinitesimally small amount of time!) because it takes no time to move over a *point*. If this surprises you, look at your answers to parts *a* and *b*: To move from +1.0 m to +4.0 m took 1.7 s; from +2.0 m to +3.0 m it took 0.6 s. As we make the interval smaller and smaller, the time it takes to move over that interval gets smaller. If we consider just the point at +2.5 m, the distance shrinks to zero, and so does the time it takes to cross the point.

2.7 (*a*) Normal: $3.50\text{ m} - 1.00\text{ m} = 2.50$ m; slow: $2.25\text{ m} - 1.00\text{ m} = 1.25$ m (half the distance, as one would expect). (*b*) Normal: $(2.50\text{ m})/(1.50\text{ s}) = 1.67$ m/s; slow: $(1.25\text{ m})/(1.50\text{ s}) = 0.83$ m/s (note units of meter per second). (*c*) Normal: $1.50\text{ s} - 0.60\text{ s} = 0.90$ s; slow: $3.00\text{ s} - 1.20\text{ s} = 1.80$ s (twice as long, as you expect). (*d*) Normal: $(1.50\text{ m})/(0.90\text{ s}) = 1.67$ m/s; slow: $(1.50\text{ m})/(1.80\text{ s}) = 0.83$ m/s (same as the values you calculated in part *b*).

2.8 (*a*) At $t = 1.3$ s. (*b*) $(3.4\text{ m} - 1.0\text{ m})/(1.3\text{ s}) = 1.8$ m/s.

2.9 (*a*) The average speed is

$$\frac{\text{distance traveled}}{\text{time taken}} = \frac{0.5\text{ km}}{6.0\text{ min}} \times \frac{1000\text{ m}}{1.0\text{ km}} \times \frac{1.0\text{ min}}{60\text{ s}} = 1.4\text{ m/s}.$$

(*b*) The average speed is

$$\frac{\text{distance traveled}}{\text{time taken}} = \frac{0.5\text{ km}}{2.0\text{ min}} \times \frac{1000\text{ m}}{1.0\text{ km}} \times \frac{1.0\text{ min}}{60\text{ s}} = 4.2\text{ m/s}.$$

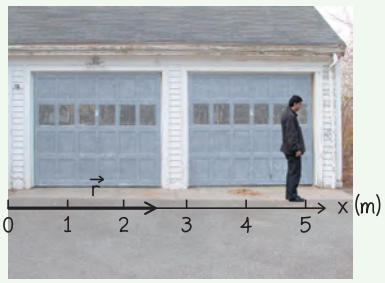
(*c*) $(1.4\text{ m/s} + 4.2\text{ m/s})/2 = 2.8$ m/s. This is not equal to the answer worked out in Example 2.3 because average speed is obtained not by averaging speeds over different segments of a trip but by dividing the distance traveled by the duration of the entire trip.

2.10 (*a*) Negative. The *x* component of the displacement is $-2.3\text{ m} - (-1.2\text{ m}) = -1.1$ m. Because the *x* component of the displacement is negative, the *x* component of the average velocity is negative. (*b*) Negative again. The *x* component of the displacement is $-1.2\text{ m} - (+1.2\text{ m}) = -2.4$ m, and so the *x* component of the average velocity is negative as well. (*c*) Yes. In both cases, the object moves in the direction of decreasing *x* and so must have a negative *x* component of the average velocity.

2.11 (*i*) and (*iv*) are scalars because no direction is required in order to specify these quantities completely; (*ii*) and (*iii*) are vectors because direction is required to specify them completely.

2.12 (*a*) Because my feet in Figure 2.18 are at $x = +4.8$ m, the *x* coordinate of the vector pointing from the origin to my feet is $x = +4.8$ m. (*b*) The vector $\vec{r}$ is obtained by multiplying the unit vector by +4.8 m: $\vec{r} = (+4.8\text{ m})\vec{\imath}$. (*c*) It represents my position in the frame shown in Figure 2.18. (*d*) Yes. It is a vector that is 2.5 m long and points to the right (Figure S2.12). This vector could represent my position at an earlier instant, or it could be a vector representing the position of the right edge of the left garage door.

Figure S2 .12

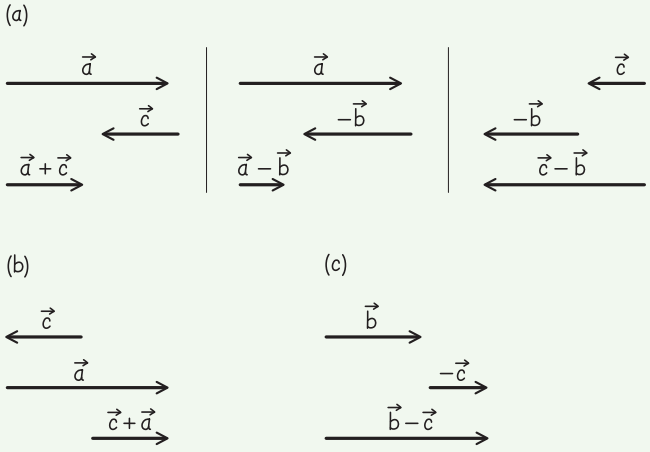


2.13 (*a*) $\Delta x = x_f - x_i = +4.8\text{ m} - (+2.2\text{ m}) = +2.6$ m. (*b*) $\Delta\vec{r} = (+2.6\text{ m})\vec{\imath}$.

2.14 (*a*) Positive. If the object is initially at $x_i = -2$ m and moves to $x_f = -1$ m, then the *x* component of the displacement is $\Delta x = -1\text{ m} - (-2\text{ m}) = +1$ m. (*b*) No. Imagine any case where the object moves from a negative value of *x* to a positive value, and the displacement is again positive. Whenever positive *x* values are to the right of the origin, the *x* component of any displacement to the right is positive, regardless of the initial and final locations.

2.15 (*a*) See Figure S2.15*a*. (*b*) Yes, because addition is commutative: $\vec{a} + \vec{c} = \vec{c} + \vec{a}$; see Figure S2.15*b*. (*c*) No, because subtraction is not commutative: $\vec{a} - \vec{c} \neq \vec{c} - \vec{a}$; see Figure S2.15*c*.

Figure S2.15



2.16 (*a*) The time interval between frames 6 and 17 is $\Delta t = t_f - t_i = 5.33\text{ s} - 1.67\text{ s} = 3.66$ s; the distance traveled is 2.0 m (from $x = +3.8$ m to $x = +4.8$ m and back). Average speed: $(2.0\text{ m})/(3.66\text{ s}) = 0.55$ m/s; *x* component of average velocity: $\Delta x = x_f - x_i = +3.8\text{ m} - (+3.8\text{ m}) = 0$, and so

$$v_{x,\text{av}} = \frac{\Delta x}{\Delta t} = \frac{0}{\Delta t} = 0.$$

Average velocity: $\vec{v}_{\text{av}} = v_{x,\text{av}}\vec{\imath} = (0)\vec{\imath} = \vec{0}$.

(*b*) The time interval between frames 1 and 17 is $\Delta t = t_f - t_i = 5.33\text{ s} - 0 = 5.33$ s; the distance traveled is 4.8 m (3.8 m from $x = +1.0$ m to $x = +4.8$ m and then 1.0 m back to $x = +3.8$). Average speed: $(4.8\text{ m})/(5.33\text{ s}) = 0.90$ m/s; *x* component of average velocity: $\Delta x = x_f - x_i = +3.8\text{ m} - (+1.0\text{ m}) = +2.8$ m, and so

$$v_{x,\text{av}} = \frac{\Delta x}{\Delta t} = \frac{+2.8\text{ m}}{5.33\text{ s}} = +0.53\text{ m/s}.$$

Average velocity: $\vec{v}_{\text{av}} = v_{x,\text{av}}\vec{\imath} = (+0.53\text{ m/s})\vec{\imath}$.

2.17 From Figure 2.30 you see that, at $t = 2.0$ s, the object is at $x = +2.8$ m. Substituting $t_i = 2.0$ s, $x_i = +2.8$ m, and $v_x = +0.60$ m/s into Eq. 2.19 yields

$$\begin{aligned} x_f &= +2.8 \text{ m} + (+0.60 \text{ m/s})(t_f - 2.0 \text{ s}) \\ &= +2.8 \text{ m} + (+0.60 \text{ m/s})t_f - (+0.60 \text{ m/s})(2.0 \text{ s}) \\ &= +2.8 \text{ m} + (+0.60 \text{ m/s})t_f - 1.2 \text{ m} \\ &= +1.6 \text{ m} + (+0.60 \text{ m/s})t_f, \end{aligned}$$

which is identical to the result in Example 2.10.

2.18 (*a*) $x_2 = 0$, $x_9 = +0.314$ m;

$$t_2 = 2 \text{ intervals} \times \frac{0.0300 \text{ s}}{\text{interval}} = 0.0600 \text{ s}$$

$$t_9 = 9 \text{ intervals} \times \frac{0.0300 \text{ s}}{\text{interval}} = 0.2700 \text{ s}$$

$$v_{x,\text{av}} = \frac{\Delta x}{\Delta t} = \frac{+0.314 \text{ m}}{0.210 \text{ s}} = +1.50 \text{ m/s}.$$

(*b*) $x_8 = +0.250$ m, $t_8 = 0.240$ s, so

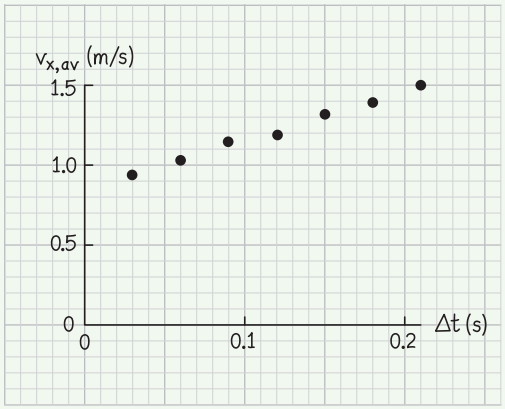
$$v_{x,\text{av}} = \frac{\Delta x}{\Delta t} = \frac{+0.250 \text{ m}}{0.180 \text{ s}} = +1.39 \text{ m/s}.$$

(*c*) See Table S2.18. (*d*) Decreases. Because the ball speeds up as it falls, averaging over a shorter time interval involves lower velocities, and therefore the average decreases (see also Figure 2.32). (*e*) See Figure S2.18. (*f*) Note that extrapolating the data to $\Delta t = 0$ does not yield a zero average velocity. This makes sense: The ball's velocity is nonzero before and after passing position 2, and so neither the average value nor the instantaneous value can be zero at position 2.

Table S2.18 Average velocities

Position	Δx (m)	Δt (s)	$\Delta x/\Delta t$ (m/s)
7	+0.197	0.1500	+1.31
6	+0.143	0.1200	+1.19
5	+0.102	0.0900	+1.13
4	+0.061	0.0600	+1.02
3	+0.028	0.0300	+0.93
2	0	0	–

Figure S2.18



2.19 (*a*) Because the derivative with respect to time of a constant is zero, you have, from Eq. 2.22, $v_x = dx/dt = 0$. If the position is constant, then its x coordinate is also constant, and so Δx is always zero. Therefore $\Delta x/\Delta t = 0$ for any finite time interval Δt. Substituting $\Delta x/\Delta t = 0$ into Eq. 2.21 yields $v_x = 0$ because the limit of zero is zero. (*b*) If $x = ct$, then $v_x = dx/dt = d(ct)/dt = c\,dt/dt = c$. Over a finite time interval Δt, you have $x_i = ct_i$ and $x_f = c(t_i + \Delta t)$. Therefore $\Delta x = c(t_i + \Delta t) - ct_i = c\Delta t$, and so $\Delta x/\Delta t = c$. Because c does not depend on Δt, the limit in Eq. 2.21 has no effect and so $v_x = c$. This result makes sense because the relationship $x = ct$ means x changes by a constant amount c each second. Therefore the object travels the same distance in equal time intervals, and so its velocity must be constant.

Chapter 3

3.1 (*a*) The x component of the average acceleration is the change in the x component of the velocity divided by the time interval during which that change took place:

$$\frac{+5.0 \text{ m/s} - 0}{1.0 \text{ s}} = +5.0 \text{ m/s}^2.$$

(*b*)
$$\frac{+10 \text{ m/s} - (+5.0 \text{ m/s})}{2.0 \text{ s}} = +2.5 \text{ m/s}^2.$$

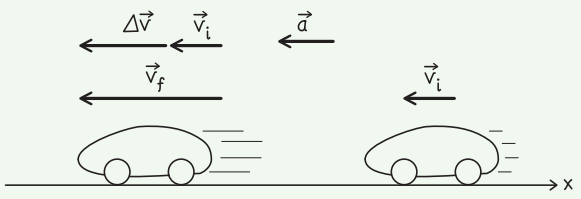
(*c*)
$$\frac{+10 \text{ m/s} - 0}{3.0 \text{ s}} = +3.3 \text{ m/s}^2.$$

3.2 (*a*) The minus signs on the two v_x values tell you that the car is traveling in the negative x direction. (*b*) $\Delta v_x = v_{x,f} - v_{x,i} = -2$ m/s $-$ $(-10$ m/s$) = +8$ m/s; because Δv_x is positive, the vector $\Delta\vec{v}$ points in the positive x direction. (*c*) $\vec{a}$ always points in the same direction as $\Delta\vec{v}$ and so points in the positive x direction. Consequently the x component of the acceleration is positive. (*d*) The acceleration points in the direction opposite the direction of the velocity, and so the car is slowing down.

3.3 (*a*) For both graphs, x decreases with time, which means both velocities are in the negative x direction, and so the x components of the velocities are negative. (*b*) The speeds are increasing in part *a* (more distance covered in equal time intervals as the $x(t)$ curve bends down; decreasing in part *b* as the curve bends up. (*c*) From part *a* you know that the velocity is always in the negative x direction; from part *b* you know that the speed is increasing in *a* and decreasing in *b*. Combining these results, you see that $\Delta\vec{v}$ points in the negative x direction in *a* and in the positive x direction in *b*, and so Δv_x is negative in *a* and positive in *b*. (*d*) The x components are negative in *a* and positive in *b* ($\vec{a}$ always points in the same direction as $\Delta\vec{v}$, and so the algebraic signs of Δv_x and a_x are the same).

3.4 (*a*) The car is moving in the negative x direction, and so $\vec{v}$ must point in that direction. To increase the car's speed, $\Delta\vec{v}$ must point in the negative x direction (Figure S3.4). Because $\vec{a}$ and $\Delta\vec{v}$ always point in the same direction, $\vec{a}$ also points in the negative x direction. (*b*) The figure shows a car slowing down in the negative x direction. Motion in the negative x direction implies a negative slope of the $x(t)$ curve, which eliminates the two graphs in Figure 3.2. Because the car slows down, it should cover increasingly smaller distances (smaller Δx for a given Δt at later instants t). This is so for Figure 3.3*b*.

Figure S3.4



3.5 When the book and the paper are held side by side, the book falls faster. When the paper is on top of the book, they fall at the same rate. The difference is due to air resistance, which slows down the light piece of paper but not the heavier book. When the paper is on top of the book, the paper is shielded from the flow of air and so feels no air resistance.

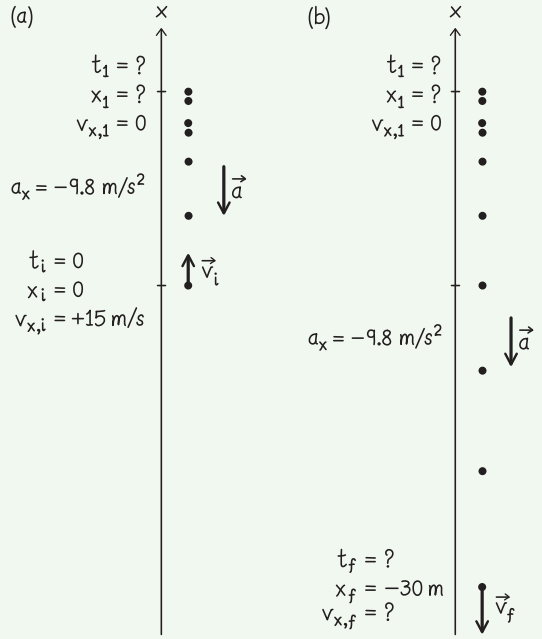
3.6 In Figure 3.6*b* you can see that Δx goes from -2.2 mm at $n = 2$ to -7.0 mm at $n = 10$: The displacement increases from one exposure to the next as n increases—the ball is accelerating. Because the magnitude of the displacement is increasing with time, the magnitude of the velocity must also be increasing.

3.7 (*a*) The speed increases because the object covers greater and greater distances in equal time intervals. (*b*) The x component of the velocity is negative because the motion is in the negative x direction. From part *a* you know that the magnitude of the velocity increases, and so the x component of the velocity becomes increasingly negative—it *decreases*. (*c*) Because the object speeds up in the negative x direction, the acceleration points in the negative x direction and so a_x is negative.

3.8 The ball's speed is about 20 m/s after 1 s and 30 m/s after 2 s. All your throwing has done is change the initial speed from zero (the initial speed if the ball is released from rest) to 10 m/s. Once the ball leaves your hand at this initial speed, the only thing affecting its motion is the acceleration due to gravity, which increases the speed at the rate of about 10 m/s^2.

3.9 See Figure S3.9. Because the rock reverses its motion, the motion diagram is split into two parts: one for the upward motion and the other for the downward motion. If you take up as the positive direction, the x component of the rock's acceleration is $a_x = -9.8$ m/s^2. With the origin chosen to be at the top of the cliff, the initial conditions are $t_i = 0$, $x_i = 0$, and $v_{x,i} = +15$ m/s. At the top of its trajectory, the rock reverses direction, and so the velocity is zero at that position. You do not know its position or the instant at which it reaches the top, and so your diagram must show $t_1 = ?$, $x_1 = ?$, and $v_{x,1} = 0$. The final conditions at the instant the rock reaches the water are $t_f = ?$, $x_f = -30$ m, and $v_{x,f} = ?$

Figure S3.9



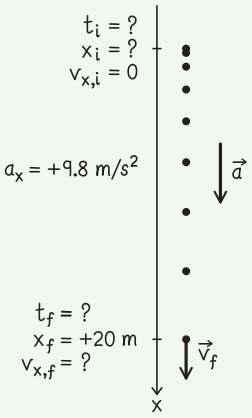
3.10 From the motion diagram in Figure 3.11 you have $t_i = 0$, $x_i = 0$, and $v_{x,i} = 0$. Substituting these values into Eq. 3.9 gives $x_f = \frac{1}{2}a_x t_f^2$ or, solving for t_f: $t_f = \sqrt{2x_f/a_x}$. Substituting the values for x_f and a_x, you get $t_f = \sqrt{2(300\text{ m})/(9.8\text{ m/s}^2)} = 7.8$ s. Equation 3.10 then yields $v_{x,f} = v_{x,i} + a_x t_f = 0 + (9.8\text{ m/s}^2)(7.8\text{ s}) = 77$ m/s. (Alternatively, you can substitute the values for a_x and Δx into Eq. 3.13 to obtain the same result.)

3.11 Figure S3.11 shows the motion diagram for the new choice of x axis. The initial conditions are as before; the final conditions are $t_f = ?$, $x_f = +h$, and $v_{x,f} = ?$. The acceleration is now positive ($a_x = +g$) because it points in the same direction as the x axis, and so Eq. 3.9 is

$$h = 0 + 0 + \tfrac{1}{2}gt_f^2,$$

which yields the same result as before for t_f. The x component of the final velocity is $v_{x,f} = +gt_f$. This x component of the velocity is equal in magnitude to the velocity in Example 3.5 but opposite in sign (because the axis has been inverted).

Figure S3.11



3.12 Applying Eq. 3.10 to the downward part of the motion, with $v_{x,i} = 0$, $a_x = -g$, and $t = v_{x,i}/g$, you get

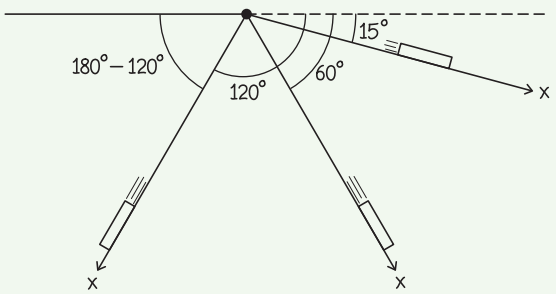
$$v_{x,f} = 0 - g\frac{v_{x,i}}{g} = -v_{x,i}.$$

The stone hits the ground at the same speed it had when it was launched.

3.13 Figure S3.13 shows that any angle $180° > \theta > 90°$ gives the same result you would get for an angle of $180° - \theta$. All that matters is the angle the incline makes with the floor; once you go beyond 90°, the cart faces the other way, but that is immaterial. (Compare, for example, the 60° angle with $180° - 60° = 120°$.) Note that Eq. 3.20 properly accounts for this fact because $\sin\theta = \sin(180° - \theta)$. For example,

$$a_{120°} = g\sin(120°) = g\sin(180° - 120°) = a_{60°}.$$

Figure S3.13



3.14 The first derivative, dropping the subscript f on t_f because it represents an (arbitrary) instant, is

$$\frac{dx_f}{dt} = \frac{d}{dt}\left(x_i + v_{x,i}t + \tfrac{1}{2}a_x t^2\right) = 0 + v_{x,i} + a_x t.$$

The expression on the right side is the x component of the velocity of an object moving at constant acceleration (Eq. 3.10). The second derivative yields the x component of the (constant) acceleration:

$$\frac{d^2x_f}{dt^2} = \frac{d}{dt}\frac{dx_f}{dt} = \frac{d}{dt}(v_{x,i} + a_x t) = 0 + a_x.$$

Chapter 4

4.1 (*a*) Yes, because each of the $v_x(t)$ curves is a straight line (constant slope = constant acceleration). (*b*) The concrete, because the magnitude of the acceleration is largest when the magnitude of Δv_x for a given time interval is largest—that is, when the slope of the line is the steepest.

4.2 (*a*) $\Delta v_{1x} = v_{1x,f} - v_{1x,i} = +0.63\ \text{m/s} - (+0.14\ \text{m/s}) = +0.49\ \text{m/s}$. (*b*) $\Delta v_{2x} = v_{2x,f} - v_{2x,i} = +0.14\ \text{m/s} - (+0.63\ \text{m/s}) = -0.49\ \text{m/s}$. (*c*) The changes are equal in magnitude but opposite in sign, meaning that the velocity of one increases and the velocity of the other decreases.

4.3 No. The magnitude of the change in the x component of the velocity of the standard cart is always half that of the half cart. If you lower $v_{sx,i}$, the only thing that happens is that the velocity-versus-time graph shrinks in the vertical direction; the x component of the velocity of the standard cart still ends up being positive. If you increase $v_{sx,i}$, the graph expands in the vertical direction and again the x component of the final velocity of the cart is positive.

4.4 $\Delta v_{ux} = v_{ux,f} - v_{ux,i} = +0.21\ \text{m/s} - (0\ \text{m/s}) = +0.21\ \text{m/s}$; $\Delta v_{sx} = v_{sx,f} - v_{sx,i} = -0.21\ \text{m/s} - (+0.42\ \text{m/s}) = -0.63\ \text{m/s}$. Therefore $\Delta v_{ux}/\Delta v_{sx} = (+0.21\ \text{m/s})/(-0.63\ \text{m/s}) = -1/3$.

4.5 $\Delta v_{mx} = v_{mx,f} - v_{mx,i} = +0.07\ \text{m/s} - (0\ \text{m/s}) = +0.07\ \text{m/s}$; $\Delta v_{px} = v_{px,f} - v_{px,i} = -0.20\ \text{m/s} - (+0.27\ \text{m/s}) = -0.47\ \text{m/s}$; $\Delta v_{px}/\Delta v_{mx} = -6.7$.

4.6 Greater. Δv_{ux} is smaller than Δv_{sx}, which tells you that the cart of unknown inertia puts up more resistance to a change in its velocity and hence has more inertia.

4.7 (*a*) Extensive. The inertia of a double cart is equal to the sum of the inertias of the individual carts. (*b*) Intensive. The velocity of a double cart is equal to the velocity of each of the individual carts. (*c*) Extensive. The product of inertia and velocity of a double cart is equal to the double inertia times the velocity of the assembly: $2m_s v_d$; the sum of the product of inertia and velocity of each individual cart is equal to two times the single inertia times the velocity of the assembly: $2m_s v_d$. In general, the product of an extensive and an intensive quantity is extensive.

4.8 (*a*) Extensive. If I separate an amount of money into two piles, then the sum of the amounts of money in each pile equals the original sum of money. (*b*) Not extensive. The sum of the temperatures in each half of a room is not equal to the temperature of the entire room. (*c*) Not extensive. The sum of the humidities in each half of a greenhouse is not equal to the humidity of the entire greenhouse. (*d*) Extensive. The volume of a bathtub is equal to the sum of the volumes of each half of the bathtub.

4.9 The x component of the change in momentum of the standard cart is positive, so the change in momentum $\Delta\vec{p}_s$ points in the positive x direction. The x component of the change in momentum of the red cart is negative and so $\Delta\vec{p}_r$ points in the negative x direction.

4.10 (*a*) Nonzero. The puck's velocity decreases, so its momentum changes. (*b*) Zero. The cart's velocity is constant. (*c*) Nonzero. The collision changes the cart's velocity. (*d*) Zero, see Eq. 4.9. (*e*) Nonzero. The cart's velocity changes.

4.11 (*a*) No. If you push with your legs against the sled, you are interacting with the sled and so you are not isolated. (*b*) Ignoring any interaction between sled and ice, however, you and the sled form an isolated system.

4.12 (*a*) The magnitude of the changes in momentum is $0.061\ \text{kg}\cdot\text{m/s}$, so the collision transfers $0.061\ \text{kg}\cdot\text{m/s}$. (*b*) Cart 1 gains that amount (Δp_{1x} is positive); cart 2 loses that same amount. So, momentum is transferred from cart 2 to cart 1.

4.13 (*a*) The magnitude of the impulse delivered to cart 1 is equal to the magnitude of the momentum change of cart 1, which we calculated in Checkpoint 4.13: $J = 0.061\ \text{kg}\cdot\text{m/s}$. (*b*) The impulse points in the same direction as the momentum change (in the positive x direction), so $\vec{J} = (+0.061\ \text{kg}\cdot\text{m/s})\hat{i}$. (*c*) No. While momentum is conserved, the momentum of a system or object can change due to transfer of momentum.

4.14 (*a*) No, because the inertias are not equal: $\Delta v_{1x} = +0.17\ \text{m/s} - 0 = +0.17\ \text{m/s}$; $\Delta v_{2x} = -0.17\ \text{m/s} - (+0.34\ \text{m/s}) = -0.51\ \text{m/s}$. (*b*) $m_1/m_2 = (0.36\ \text{kg})/(0.12\ \text{kg}) = 3$; $\Delta v_{2x}/\Delta v_{1x} = (-0.51\ \text{m/s})/(0.17\ \text{m/s}) = -3$. (*c*) $p_{1x,i} = m_1 v_{1x,i} = (0.36\ \text{kg})(0) = 0$; $p_{2x,i} = (0.12\ \text{kg})(+0.34\ \text{m/s}) = +0.041\ \text{kg}\cdot\text{m/s}$; $p_{1x,f} = (0.36\ \text{kg})(+0.17\ \text{m/s}) = +0.061\ \text{kg}\cdot\text{m/s}$; $p_{2x,f} = (0.12\ \text{kg})(-0.17\ \text{m/s}) = -0.020\ \text{kg}\cdot\text{m/s}$. (*d*) $p_{1x,i} + p_{2x,i} = 0\ \text{kg}\cdot\text{m/s} + (+0.041\ \text{kg}\cdot\text{m/s}) = +0.041\ \text{kg}\cdot\text{m/s}$. (*e*) $p_{1x,f} + p_{2x,f} = +0.061\ \text{kg}\cdot\text{m/s} + (-0.020\ \text{kg}\cdot\text{m/s}) = +0.041\ \text{kg}\cdot\text{m/s}$. (*f*) Yes. $\Delta p_{1x} = +0.061\ \text{kg}\cdot\text{m/s} - 0 = +0.061\ \text{kg}\cdot\text{m/s}$; $\Delta p_{2x} = -0.020\ \text{kg}\cdot\text{m/s} - (+0.041\ \text{kg}\cdot\text{m/s}) = -0.061\ \text{kg}\cdot\text{m/s}$. This is so because the momentum of the system does not change.

4.15 (*a*) Yes. Just as we do for the carts in Figure 4.3, we can ignore the interaction between the carts and the low-friction track, and so the two carts constitute an isolated system. (*b*) For an isolated system, the momentum law tells us that the momentum of the system cannot change.

4.16 There are two ways to solve this problem. The hard way is to apply the law of conservation of momentum to each collision separately. The easy way is to look at the momentum of the system at only two instants: before the first collision and after the final one. The system's momentum before the first collision is $mv_{x,i}$, where m is the inertia of a single railroad car. After all the cars collide, they move together at the same final speed $v_{x,f}$. The system's final momentum is therefore $4mv_{x,f}$. Because momentum is conserved, the final momentum must equal the initial momentum. Therefore $4mv_{x,f} = mv_{x,i}$, and thus $v_{x,f} = \frac{1}{4}v_{x,i}$. The speed of the four-car train is one-quarter the initial speed of the first car.

Chapter 5

5.1 No. The relative speed $|\vec{v}_2 - \vec{v}_1|$ takes into account the directions of $\vec{v}_1$ and $\vec{v}_2$, but the difference in speeds $v_2 - v_1$ doesn't. For example, if $v_{1x} = +2\ \text{m/s}$ and $v_{2x} = -2\ \text{m/s}$, then $v_{12} = |v_{2x} - v_{1x}| = 4\ \text{m/s}$ but $v_2 - v_1 = 2\ \text{m/s} - 2\ \text{m/s} = 0$.

5.2 (*a*) Totally inelastic, because after the collision the relative speed is zero. (*b*) Elastic, because there is no change in the relative speed. (*c*) Whether or not momentum is constant depends only on whether the system is isolated, not on the type of collision. In part *a*, the system (ball + glove) is not isolated because the glove interacts with something outside the system (the outfielder). Thus the sum of the momenta is not constant. In part *b*, the system (two balls) is isolated because, if we ignore friction, neither ball interacts

with any object external to the system. Therefore the sum of the momenta is constant.

5.3 Yes. Consider an object that is moving at speed v and consists of two pieces having inertias m_1 and m_2. The kinetic energy of the object, $\frac{1}{2}(m_1 + m_2)v^2$, is equal to the sum of the kinetic energies of its two parts, $\frac{1}{2}m_1v^2 + \frac{1}{2}m_2v^2$.

5.4 Denote the x component of the initial velocity of the moving cart $v_{x,\mathrm{i}}$ and the x component of the final velocity of the combined carts $v_{x,\mathrm{f}}$. Conservation of momentum requires that the two carts have a final momentum equal to the initial momentum, and so $mv_{x,\mathrm{i}} = 2mv_{x,\mathrm{f}}$. Thus $v_{x,\mathrm{f}} = \frac{1}{2}v_{x,\mathrm{i}}$. The initial kinetic energy is $\frac{1}{2}mv_{x,\mathrm{i}}^2$; the final kinetic energy is $\frac{1}{2}(2m)v_{x,\mathrm{f}}^2 = m\left(\frac{1}{2}v_{x,\mathrm{i}}\right)^2 = \frac{1}{4}mv_{x,\mathrm{i}}^2$. This is half the initial kinetic energy.

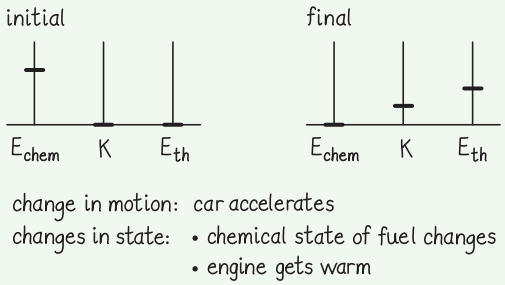
5.5 Increase. The system has some kinetic energy before the collision but zero kinetic energy after. This decrease in kinetic energy is accompanied by a change in the state of the dough (it changes shape). This change in state represents a change in the internal energy of the dough. If the sum of the internal and kinetic energies of the system is to remain constant, the internal energy must increase.

5.6 You could throw something against the cart, push it with your hand, blow against it, or ignite an explosive next to it. Changes in state or motion: The motion of the thrown object changes when it collides with the cart. To push or blow, you have to move your muscles, which causes chemical reactions that change your chemical state. The ignition causes the chemical state of the explosive to change.

5.7 (*a*) No. The momentum of the system is zero before the spring expands and nonzero afterward. (*b*) No, because $\Delta p \neq 0$. There is an external interaction with the post that holds the left end of the spring fixed. Without this interaction, the expanding spring would not cause the cart to move.

5.8 See Figure S5.8. The car engine heats up, and so part of the chemical energy is converted to thermal energy. The sum of the changes ΔK and ΔE_{th} is equal in magnitude to the change ΔE_{chem}.

Figure S5.8



5.9 (*a*) No. Changing the magnitude of an object's momentum means changing the object's speed, and this automatically means changing its kinetic energy. (*b*) Yes. If the change in kinetic energy is compensated by a change in the internal energy of the object, its energy remains unchanged. (*c*) Yes, for both parts. With a system of objects, it is possible to keep the kinetic energy constant while changing the magnitude of the momentum. Imagine two identical carts on a low-friction track moving away from each other in opposite directions at speed v. The momentum of the system is zero, and the kinetic energy of the system is mv^2. If the velocity of one of the two carts is reversed by some interaction, the system's kinetic energy remains unchanged but the momentum of the system becomes $2mv$. For the energy of the system, use the same reasoning you used to answer part *b*.

5.10 For a single object, if mv_x and mv_x^2 remain constant, then the product of these two quantities must also remain constant: $(mv_x)(mv_x^2) = m^2v_x^3 =$ constant. Because m does not change for a single, isolated object, $m^2v_x^3/m = mv_x^3$ must also be constant. Therefore mv_x^3 indeed remains constant for a single object.

In an elastic collision between two objects of identical inertia m with velocities v_{1x} and v_{2x}, the system's momentum, $mv_{1x} + mv_{2x}$, remains constant. The system's kinetic energy is also constant, and so $mv_{1x}^2 + mv_{2x}^2$ is constant. The product of these two must also be constant: $m^2v_{1x}^3 + m^2v_{1x}^2v_{2x} + m^2v_{1x}v_{2x}^2 + m^2v_{2x}^3 = m(mv_{1x}^3 + mv_{2x}^3) + mv_{1x}v_{2x}(mv_{1x} + mv_{2x}) =$ constant. From this we see that $(mv_{1x}^3 + mv_{2x}^3)$ is constant if the second term, $mv_{1x}v_{2x}(mv_{1x} + mv_{2x})$, is constant. The part in parentheses is the system's momentum and is constant. The inertia m is also constant, so all we need to know is whether or not $v_{1x}v_{2x}$ is constant. Suppose $v_{1x,\mathrm{i}}$ changes by Δv_{1x}. Conservation of momentum then requires $v_{2x,\mathrm{i}}$ to change by $-\Delta v_{1x}$. So the product of the final velocities becomes $(v_{1x,\mathrm{i}} + \Delta v_{1x})(v_{2x,\mathrm{i}} - \Delta v_{1x}) = v_{1x,\mathrm{i}}v_{2x,\mathrm{i}} + \Delta v_{1x}(v_{2x,\mathrm{i}} - v_{1x,\mathrm{i}}) - (\Delta v_{1x})^2$. This can be equal to the product of the initial velocities, $v_{1x,\mathrm{i}}v_{2x,\mathrm{i}}$, only if the change in velocity Δv_{1x} is equal to $v_{2x,\mathrm{i}} - v_{1x,\mathrm{i}}$. This can indeed happen but, in general, need *not* be so. In general, therefore, mv_x^3 is not unchanged in an elastic collision.

5.11 No. If one object is initially at rest, then the other object must be moving in order for the collision to occur, and so the initial momentum of the system is nonzero. Because the system is isolated, its momentum is unchanged, and so the system's final momentum, too, is nonzero. Therefore the combination of the two objects has a nonzero velocity after the collision, which means the kinetic energy cannot be zero.

5.12 From Eq. 5.12 you know that $v = \sqrt{2K/m}$. Substituting 1.2×10^8 J for the kinetic energy K and 1200 kg for the inertia m gives $v = 4.5 \times 10^2$ m/s, which is more than 1000 mph! (Apparently it is not possible to convert all of the energy in gasoline to kinetic energy! You'll see why in Chapter 21.)

5.13 No. Cart 1 gets twice as much energy as cart 2: $K_{1\mathrm{f}} = \frac{1}{2}m_1v_{1\mathrm{f}}^2 = \frac{1}{2}(0.25\ \mathrm{kg})(2.0\ \mathrm{m/s})^2 = 0.50$ J, and $K_{2\mathrm{f}} = \frac{1}{2}m_2v_{2\mathrm{f}}^2 = \frac{1}{2}(0.50\ \mathrm{kg}) \cdot (1.0\ \mathrm{m/s})^2 = 0.25$ J. The reason is that the system's final momentum needs to be zero, and so $v_{1\mathrm{f}}$ must be $2v_{2\mathrm{f}}$. Because $m_2 = 2m_1$, you have $K_{2\mathrm{f}} = \frac{1}{2}m_2v_{2\mathrm{f}}^2 = \frac{1}{2}(2m_1)\left(\frac{1}{2}v_{1\mathrm{f}}\right)^2 = \frac{1}{4}m_1v_{1\mathrm{f}}^2 = \frac{1}{2}K_{1\mathrm{f}}$.

Chapter 6

6.1 Because $\vec{v}_{\mathrm{Eo}} = \vec{v}_{\mathrm{EM}} + \vec{v}_{\mathrm{Mo}}$, you can write $v_{\mathrm{Mo}x} = v_{\mathrm{Eo}x} - v_{\mathrm{EM}x}$. With $v_{\mathrm{EM}x} = -3.0$ mm/frame, you have $v_{\mathrm{M1}x} = 0 - (-3.0\ \mathrm{mm/frame}) = +3.0$ mm/frame and $v_{\mathrm{M2}x} = (3.6\ \mathrm{mm/frame}) - (-3.0\ \mathrm{mm/frame}) = +6.6$ mm/frame.

6.2 (*a*) There is no change in the state of either cart, and so the internal energy of either cart is not changing. From observer E's point of view, both velocities are constant and so both kinetic energies are constant. Observer M reaches the same conclusion: Both carts are moving at constant velocity and so their kinetic energies are constant. The observers agree that the energy of each cart is constant. (*b*) and (*c*) No state change or motion change in the environment surrounding each isolated system causes any state change or motion change inside the system, and so each system is closed.

6.3 (*a*) There is no change in the state of either cart and so no change in the internal energy. For observer E, cart 1 accelerates, and so its kinetic energy K and energy E increase; the velocity of cart 2 is constant, and so its kinetic energy K and energy E are constant. For observer M, cart 1 is at rest, and so its kinetic energy K and energy E are constant; cart 2 accelerates, and so its kinetic energy K and energy E are increasing. (*b*) For observer E, a change in state in the spring is responsible for the cart's acceleration, and so the system is not closed. For observer M, changes in both state and motion occur in the environment, but the expansion of the spring cannot account for the change in the kinetic energy of the entire environment (the universe minus cart 1), which means that the energy changes don't add to zero and so the system is

not closed. (*c*) For observer E, there are no changes in state or motion in the environment that are responsible for any changes in state or motion inside the system, and so the system is closed. For observer M, there are no changes in state or motion in the environment, and so it is closed. (*d*) For observer E, the cart 1 system is not closed and its energy increases; the cart 2 system is closed, and its energy is constant. The conservation of energy law holds for both systems. For observer M, the cart 1 system is not closed, and its energy remains constant; the cart 2 system is closed, and its energy is increasing. Neither observation agrees with the law of conservation of energy.

6.4 (*a*) $v_{\mathrm{Ec}x,\mathrm{i}} = +0.40$ m/s, $v_{\mathrm{Ec}x,\mathrm{f}} = +0.80$ m/s, $\Delta K = K_{\mathrm{Ec,f}} - K_{\mathrm{Ec,i}} = \frac{1}{2}(0.12\ \mathrm{kg})(0.80\ \mathrm{m/s})^2 - \frac{1}{2}(0.12\ \mathrm{kg})(0.40\ \mathrm{m/s})^2 = 29$ mJ.
(*b*) $v_{\mathrm{Mc}x,\mathrm{i}} = -0.20$ m/s, $v_{\mathrm{Mc}x,\mathrm{f}} = +0.20$ m/s,
$\Delta K = \frac{1}{2}(0.12\ \mathrm{kg})(0.20\ \mathrm{m/s})^2 - \frac{1}{2}(0.12\ \mathrm{kg})(-0.20\ \mathrm{m/s})^2 = 0$.
(*c*) $v_{\mathrm{Mc}x,\mathrm{i}} = -0.40$ m/s, $v_{\mathrm{Mc}x,\mathrm{f}} = 0$,
$\Delta K = 0 - \frac{1}{2}(0.12\ \mathrm{kg})(-0.40\ \mathrm{m/s})^2 = 10$ mJ.

6.5 Because the momentum of the isolated system is not changing: $m_1 v_{\mathrm{E1}x,\mathrm{i}} + m_2 v_{\mathrm{E2}x,\mathrm{i}} = (m_1 + m_2) v_{\mathrm{E}x,\mathrm{f}}$, and so $v_{\mathrm{E}x,\mathrm{f}} = (m_1 v_{\mathrm{E1}x,\mathrm{i}} + m_2 v_{\mathrm{E2}x,\mathrm{i}})/(m_1 + m_2) = +0.20$ m/s. In the moving reference frame, $v_{\mathrm{M}x,\mathrm{f}} = +0.40$ m/s. See Table S6.5.

Table S6.5 Conversion of kinetic energy

Cart	Inertia (kg)	v_x (m/s) before	v_x (m/s) after	Δv_x	K (10^{-3} J) before	K (10^{-3} J) after	ΔK
Earth reference frame							
1	0.36	0	+0.20	+0.20	0	7.2	+7.2
2	0.12	+0.80	+0.20	−0.60	38.4	2.4	−36
						$\Delta K =$	−29
Reference frame moving at −0.20 m/s relative to Earth							
1	0.36	+0.20	+0.40	+0.20	7.2	28.8	+21.6
2	0.12	+1.0	+0.40	−0.60	60	9.6	−50.4
						$\Delta K =$	−29

6.6 Because the laws of the universe are the same in all inertial reference frames, the coefficient of restitution must also be the same. Another way of seeing this: Relative velocities do not depend on the velocity of the reference frame (Figure 6.11), and so $e = v_{12\mathrm{f}}/v_{12\mathrm{i}}$ is also independent of the reference frame.

6.7 Less. In the Earth reference frame, $K_{\mathrm{E,i}} = 0 + \frac{1}{2}(0.12\ \mathrm{kg}) \cdot (0.80\ \mathrm{m/s})^2 = 38$ mJ. In the zero-momentum reference frame, $K_{\mathrm{Z,i}} = \frac{1}{2}(0.36\ \mathrm{kg})(0.20\ \mathrm{m/s})^2 + \frac{1}{2}(0.12\ \mathrm{kg})(0.60\ \mathrm{m/s})^2 = 29$ mJ.

6.8 (*a*) If the jogger stays in place, he must be moving at the same speed as the belt but in the opposite direction. Therefore $v_{\mathrm{Bj}x} = -2.0$ m/s, and so $r_{\mathrm{Bj}x} = v_{\mathrm{Bj}x}\, t = (-2.0\ \mathrm{m/s})(10\ \mathrm{s}) = -20$ m. (*b*) Equation 6.3 yields $r_{\mathrm{Ej}x} = r_{\mathrm{EB}x} + r_{\mathrm{Bj}x}$. The position of the belt relative to Earth is $r_{\mathrm{EB}x} = v_{\mathrm{EB}x}\Delta t$, and the position of the jogger relative to the belt is $r_{\mathrm{Bj}x} = v_{\mathrm{Bj}x}\Delta t$. Because $v_{\mathrm{Bj}x} = -v_{\mathrm{EB}x}$, you find that $r_{\mathrm{Ej}x} = r_{\mathrm{EB}x} + r_{\mathrm{Bj}x} = v_{\mathrm{EB}x}\Delta t + v_{\mathrm{Bj}x}\Delta t = 0$.

6.9 If you define northward as the positive x direction, you have $v_{\mathrm{TP}x} = +1.2$ m/s, $v_{\mathrm{ET}x} = +3.1$ m/s, and $v_{\mathrm{Ps}x} = -0.5$ m/s. For the velocity of the spider relative to Earth, subscript cancellation gives $v_{\mathrm{Es}x} = v_{\mathrm{EP}x} + v_{\mathrm{Ps}x}$. For the velocity of the passenger relative to Earth, $v_{\mathrm{EP}x} = v_{\mathrm{ET}x} + v_{\mathrm{TP}x}$, and so $v_{\mathrm{Es}x} = (v_{\mathrm{ET}x} + v_{\mathrm{TP}x}) + v_{\mathrm{Ps}x} = (+3.1\ \mathrm{m/s}) + (+1.2\ \mathrm{m/s}) + (-0.5\ \mathrm{m/s}) = +3.8$ m/s. An observer standing by the side of the tracks sees the spider moving at 3.8 m/s northward.

6.10 (*a*) Substituting $m_1 = 3m_2 = 3m$ into Eq. 1 of Example 6.7 yields $x_{\mathrm{A}\mathrm{cm}} = (3mx_{\mathrm{A1}} + mx_{\mathrm{A2}})/(3m + m) = (3x_{\mathrm{A1}} + x_{\mathrm{A2}})/4$, one-quarter of the way from x_{A1} to x_{A2}. (*b*) $x_{\mathrm{A}\mathrm{cm}} = (m_1 x_{\mathrm{A1}} + m_2 x_{\mathrm{A2}} + m_3 x_{\mathrm{A3}})/(m_1 + m_2 + m_3)$. Substituting $m_1 = m_3 = 3m_2 = 3m$ and $x_{\mathrm{A}\mathrm{cm}} = x_{\mathrm{A2}}$ gives $x_{\mathrm{A2}} = (3mx_{\mathrm{A1}} + mx_{\mathrm{A2}} + 3mx_{\mathrm{A3}})/(3m + m + 3m) = (3x_{\mathrm{A1}} + x_{\mathrm{A2}} + 3x_{\mathrm{A3}})/7$. Then $x_{\mathrm{A3}} = 2x_{\mathrm{A2}} - x_{\mathrm{A1}}$.

6.11 (*a*) Because $m_1 = 3m_2$, Eq. 6.26 gives: before the collision $v_{\mathrm{cm}x} = [(m_1)(0) + (m_2)(+0.80\ \mathrm{m/s})]/(m_1 + m_2) = +0.20$ m/s; after the collision $v_{\mathrm{cm}x} = [(3m_2)(+0.40\ \mathrm{m/s}) + m_2(-0.40\ \mathrm{m/s})]/(3m_2 + m_2) = +0.20$ m/s. This is the x component of the velocity where the two $v_x(t)$ curves intersect. (*b*) Yes. When two objects collide, the object that initially has greater velocity always ends up with less velocity after the collision (and vice versa). Therefore there always has to be a location at which the two objects have the same velocity. At that location $v_{1x} = v_{2x} = v_x$, and so

$$v_{\mathrm{cm}x} = (m_1 v_x + m_2 v_x)/(m_1 + m_2)$$
$$= v_x(m_1 + m_2)/(m_1 + m_2) = v_x.$$

(*c*) From Eqs. 6.23 and 6.26, $v_{\mathrm{EZ}x} = v_{\mathrm{cm}x} = +0.20$ m/s. From Eq. 6.8, $v_{\mathrm{E1}x,\mathrm{i}} = v_{\mathrm{EZ}x} + v_{\mathrm{Z1}x,\mathrm{i}}$, and so $v_{\mathrm{Z1}x,\mathrm{i}} = v_{\mathrm{E1}x,\mathrm{i}} - v_{\mathrm{EZ}x} = 0 - (+0.20\ \mathrm{m/s}) = -0.20$ m/s, $v_{\mathrm{Z2}x,\mathrm{i}} = +0.6$ m/s, $v_{\mathrm{Z1}x,\mathrm{f}} = +0.2$ m/s, and $v_{\mathrm{Z2}x,\mathrm{f}} = -0.6$ m/s. Then $p_{\mathrm{Z}x,\mathrm{i}} = (3m_2)(-0.20\ \mathrm{m/s}) + (m_2)(+0.60\ \mathrm{m/s}) = 0$ and $p_{\mathrm{Z}x,\mathrm{f}} = (3m_2)(+0.20\ \mathrm{m/s}) + (m_2) \cdot (-0.60\ \mathrm{m/s}) = 0$.

6.12

$$K_{\mathrm{conv}} = \left(\tfrac{1}{2}m_1 v_1^2 + \tfrac{1}{2}m_2 v_2^2\right) - \tfrac{1}{2}(m_1 + m_2)v_{\mathrm{cm}}^2$$
$$= \left(\tfrac{1}{2}m_1 v_1^2 + \tfrac{1}{2}m_2 v_2^2\right) - \frac{1}{2}\frac{(m_1 v_{1x} + m_2 v_{2x})^2}{m_1 + m_2}$$
$$= \left(\tfrac{1}{2}m_1 v_1^2 + \tfrac{1}{2}m_2 v_2^2\right) - \frac{1}{2}\frac{(m_1^2 v_1^2 + 2m_1 m_2 v_{1x} v_{2x} + m_2^2 v_2^2)}{m_1 + m_2}$$
$$= \frac{1}{2}\frac{(m_1 v_1^2 + m_2 v_2^2)(m_1 + m_2) - (m_1^2 v_1^2 + 2m_1 m_2 v_{1x} v_{2x} + m_2^2 v_2^2)}{m_1 + m_2}$$
$$= \frac{1}{2}\frac{(m_1 m_2 v_1^2 + m_1 m_2 v_2^2) - 2m_1 m_2 v_{1x} v_{2x}}{m_1 + m_2}$$
$$= \frac{1}{2}\frac{m_1 m_2 (v_1^2 - 2v_{1x}v_{2x} + v_2^2)}{m_1 + m_2} = \frac{1}{2}\left(\frac{m_1 m_2}{m_1 + m_2}\right)(v_{2x} - v_{1x})^2$$
$$= \frac{1}{2}\left(\frac{m_1 m_2}{m_1 + m_2}\right)v_{12}^2$$

6.13 (*a*) Because $\mu = m(0.5m)/(m + 0.5m) = m/3$, $K_{\mathrm{conv}} = \frac{1}{2}(m/3)\, v^2$. Because $K_{\mathrm{i}} = \frac{1}{2}mv^2$, the fraction that is convertible is $K_{\mathrm{conv}}/K_{\mathrm{i}} = 1/3$. (*b*) If it were converted, the momentum of the isolated two-object system would change.

6.14 (*a*) $v_{12\mathrm{i}} = |0 - (+4.0\ \mathrm{m/s})| = 4.0$ m/s and $v_{12\mathrm{f}} = |(+1.5\ \mathrm{m/s}) - (-0.50\ \mathrm{m/s})| = 2.0$ m/s, so $e = (2.0\ \mathrm{m/s})/(4.0\ \mathrm{m/s}) = 0.50$.
(*b*) See Table S6.14.

Table S6.14 Conversion of kinetic energy

Object	inertia (kg)	Velocity (m/s) before	Velocity (m/s) after	Δv_x	Kinetic energy (J) before	Kinetic energy (J) after	ΔK
Earth reference frame							
1	1.0	+4.0	−0.50	−4.5	8.0	0.13	−7.9
2	3.0	0	+1.5	+1.5	0	3.4	+3.4
						$\Delta K =$	−4.5
Reference frame moving at −1.0 m/s relative to Earth							
1	1.0	+5.0	+0.50	−4.5	12.5	0.13	−12.4
2	3.0	+1.0	+2.5	+1.5	1.5	9.4	+7.9
						$\Delta K =$	−4.5

Chapter 7

7.1 (*a*) The other object is Earth; this is an attractive interaction because it makes the ball and Earth accelerate toward each other. (*b*) Repulsive, because the interaction causes the two objects to accelerate away from each other.

7.2 (*a*) At $t = 30$ ms, the velocity-versus-time graph shows $v_{1x} = 0$ and $v_{2x} = +0.55$ m/s, so $p_{1x} = (0.12\text{ kg})(0) = 0$ and $p_{2x} = (0.24\text{ kg})(+0.55\text{ m/s}) = +0.13\text{ kg}\cdot\text{m/s}$. At $t = 60$ ms, $v_{1x} = 0.37$ m/s and $v_{2x} = 0.37$ m/s, so $p_{1x} = (0.12\text{ kg})(+0.37\text{ m/s}) = +0.044\text{ kg}\cdot\text{m/s}$ and $p_{2x} = (0.24\text{ kg})(+0.37\text{ m/s}) = +0.089\text{ kg}\cdot\text{m/s}$. At $t = 90$ ms, $v_{1x} = +0.73$ m/s and $v_{2x} = +0.18$ m/s, so $p_{1x} = +0.088\text{ kg}\cdot\text{m/s}$ and $p_{2x} = +0.043\text{ kg}\cdot\text{m/s}$. (*b*) At 30 ms, $p_x = p_{1x} + p_{2x} = 0 + (+0.13\text{ kg}\cdot\text{m/s}) = +0.13\text{ kg}\cdot\text{m/s}$. At 60 ms, $p_x = (+0.044\text{ kg}\cdot\text{m/s}) + (+0.089\text{ kg}\cdot\text{m/s}) = +0.13\text{ kg}\cdot\text{m/s}$. At 90 ms, $p_x = (+0.088\text{ kg}\cdot\text{m/s}) + (+0.043\text{ kg}\cdot\text{m/s}) = +0.13\text{ kg}\cdot\text{m/s}$.

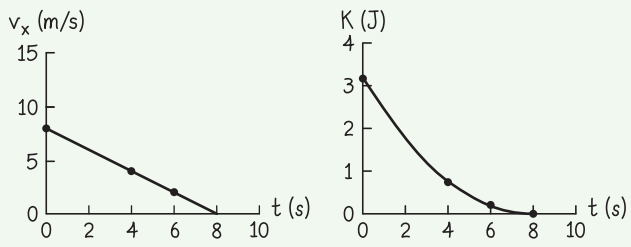
7.3 (*a*) With the velocity values from Checkpoint 7.2 and the formula $K = \frac{1}{2}mv^2$, the kinetic energies are: at $t = 30$ ms, $K_1 = 0$ and $K_2 = 0.036$ J; at $t = 60$ m/s, $K_1 = 0.0082$ J and $K_2 = 0.016$ J; at $t = 90$ ms, $K_1 = 0.032$ J and $K_2 = 0.0039$ J. (*b*) At 30 ms, $K = 0.036$ J; at 60 ms, $K = 0.025$ J; at 90 ms, $K = 0.036$ J.

7.4 (*a*) Zero, because the ball's speed is zero. (*b*) The momentum of the ball is constant only when the ball is isolated—in other words, when it is not interacting with anything. Ignoring the gravitational interaction with Earth because that interaction is much, much weaker than the ball-wall interaction, you can say that the ball is not interacting with anything before and after the collision, and so its momentum remains constant during these intervals. The ball interacts with the wall during the collision, and so now momentum is not constant for the system made up of only the ball; momentum is constant, however, for the system made up of ball plus wall.

7.5 As the spring is compressed, the kinetic energy of the cart is converted to potential energy in the deformed spring. Because the sum of the kinetic and potential energies must not change, any loss in kinetic energy must be made up by a gain in potential energy. Initially, all the energy of the system is kinetic and equal to $\frac{1}{2}mv_i^2$, where m is the inertia of the cart. In the middle drawing, the kinetic energy is zero, and so the potential energy is $\frac{1}{2}mv_i^2$.

7.6 (*a*) Let the direction in which the puck moves be the $+x$ direction. The x component of the puck's initial velocity is then $v_x = +8.0$ m/s, and the x component of its acceleration is $a_x = -1.0\text{ m/s}^2$. With that acceleration, the puck's speed is reduced by 1.0 m/s each second. So, after 4 s, the speed is half of its original value. Because kinetic energy depends on v^2, the puck's kinetic energy is reduced to one-quarter of its original value after 4 s. After 6 s, the speed is one-quarter of its original value and the kinetic energy is one-sixteenth. Based on these data, your sketches should look like Figure S7.6. (*b*) The kinetic energy is converted to thermal energy (both puck and ice get warmer).

Figure S7.6



7.7 It should be classified as incoherent configuration energy. Chemical energy is associated with changes in the arrangement of atoms in molecules and so is configuration energy. In a chemical reaction, the orientation of the molecules and the kinetic energy given to the products are random, and so chemical energy is incoherent. It is not possible to make all products move in one direction and extract (coherent) kinetic energy. (The same arguments hold for nuclear energy, the form of incoherent configuration energy that corresponds to changes in nuclear configuration.)

7.8 Both of you, because the term *conserve* means one thing when used by physicists and another thing when used by environmentalists. To physicists, energy is always conserved, which means it is neither created out of nothing nor destroyed. When you leave the lights on, electrical energy is converted to light energy and thermal energy (lamps get hot). The sum of those forms of energy is equal to the electrical energy consumed by the lights. To environmentalists, *conserving energy* means reducing or eliminating the use of source energy. When you turn off the lights, less source energy (oil, gas, coal, or nuclear fuel) needs to be converted to electrical energy. The phrase *conserving energy* has nothing to do with the law of conservation of energy, and because Earth's supply of source energy is limited, you are right to turn off the lights.

7.9 (*a*) Elastic potential energy in the spring is converted to kinetic energy in the ball and in the spring as the spring expands. The interaction is nondissipative because it can run in reverse: The ball compresses the spring. (*b*) The ball gains speed as it changes its position relative to the ground, which means potential energy is converted to kinetic energy. The interaction is nondissipative because it can run in reverse: The ball is launched up. (*c*) Kinetic energy is converted to thermal energy: The bicycle slows down because of friction, which irreversibly dissipates the kinetic energy. The interaction is dissipative because the reverse motion—the bicycle accelerates backward without the pedals moving—is not possible. (*d*) Chemical energy in the gasoline is converted to kinetic energy (and thermal energy). The conversion is irreversible and therefore the interaction is dissipative. (If you played this situation in reverse, you would see a car slowing down backward. This is possible, but close inspection would reveal details that are not possible: The driver pressing the gas pedal instead of the brake pedal and fuel being added to the tank.)

7.10 (*a*) When you throw a ball, your momentum increases in the direction opposite the direction of the ball's momentum because the momentum of the isolated system comprising you and the ball must be constant. The throwing corresponds to an explosive separation ($e > 1$). The catching corresponds to a totally inelastic collision ($e = 0$). So, when your friend catches the ball, her momentum increases in the direction in which the ball was originally moving (away from you). (Normally these momentum changes aren't noticeable because of the friction between feet and ground, which is why this checkpoint has you standing on ice.) (*b*) The thrower moves back ("recoils") in the direction opposite the direction of the thrown ball (away from the catcher). The catcher recoils in the direction in which the ball was originally moving (away from the thrower). Therefore, each throw or catch increases the momentum of each of you in the direction away from the other, meaning that the interaction is repulsive. The ball serves as the carrier—the "gauge particle"—of the interaction. (*c*) Improbable as it may seem, there actually is a mechanical model for an attractive interaction based on an exchange of particles. Imagine two people standing within arm's reach of each other and extending their arms toward each other without touching. Person A holds a ball and throws it back toward himself; person B catches it before it hits person A, and throws it back toward herself. Repeat as needed. This is perhaps not easy to implement, but the throwing back and forth does result in an attraction.

7.11 (*a*) Once the vehicles lock together, they have identical velocities, the center-of-mass velocity. If you take the x component of the car's initial velocity as positive, you get from Eq. 6.26 $v_{\text{cm}x} = [(1000\text{ kg})(+25\text{ m/s}) + (2000\text{ kg})(-25\text{ m/s})]/[1000\text{ kg} + 2000\text{ kg}] = -8.3$ m/s. Therefore $\Delta v_c = (-8.3\text{ m/s}) - (+25\text{ m/s}) = -33$ m/s and $\Delta v_v = (-8.3\text{ m/s}) - (-25\text{ m/s}) = +17$ m/s. Because it has

less inertia, the car has the larger change in velocity. Because both velocity changes occur during the same time interval, the acceleration is greater for the car, in agreement with Eq. 7.6. (*b*) $a_{cx,av} = (-33\text{ m/s})/(0.20\text{ s}) = -1.7\times10^2\text{ m/s}^2$ and $a_{cx,av} = (+17\text{ m/s})/(0.20\text{ s}) = +83\text{ m/s}^2$. The ratio is not exactly -2 because of the rounding of the velocity changes in part *a*.

7.12

$$\begin{aligned}\Delta K_E &= \tfrac{1}{2}m_E v_{E,f}^2 - \tfrac{1}{2}m_E v_{E,i}^2 = \tfrac{1}{2}m_E(v_{E,f}^2 - v_{E,i}^2)\\ &= \tfrac{1}{2}m_E[(v_{Ex,i} + \Delta v_{Ex})^2 - v_{E,i}^2]\\ &= m_E v_{E,i}\Delta v_{Ex} + \tfrac{1}{2}m_E(\Delta v_{Ex})^2.\end{aligned}$$

Rearranging terms and substituting Δp_{Ex} for $m_E\Delta v_{Ex}$, you get

$$\begin{aligned}\Delta K_E &= v_{Ex,i}(m_E\Delta v_{Ex}) + \tfrac{1}{2}\frac{(m_E\Delta v_{Ex})^2}{m_E}\\ &= v_{Ex,i}\Delta p_{Ex} + \frac{(\Delta p_{Ex})^2}{2m_E}.\end{aligned}$$

The first term on the right of the final equals sign is zero because in the reference frame of Earth, $v_{Ex,i} = 0$. The second term to the right of the final equals sign is negligibly small compared to the corresponding term for the cart because the numerators are the same (conservation of momentum requires that $\Delta p_E = -\Delta p_c$) and $m_E \gg m_c$. Therefore $\Delta K_E \approx 0$, and the only change in kinetic energy in the Earth-cart system must be that of the cart.

7.13

$$\begin{aligned}\Delta U_{12} &= U(x_2) - U(x_1)\\ \Delta U_{21} &= U(x_1) - U(x_2) = -\Delta U_{12}\\ \Delta U_{1\to2\to1} &= \Delta U_{12} + \Delta U_{21}\\ &= [U(x_2) - U(x_1)] + [U(x_1) - U(x_2)]\\ &= 0.\end{aligned}$$

The potential energy change as an object moves in one direction has a sign, and the potential energy change when the object moves back to its original position has the opposite sign and the same magnitude. Thus the change in potential energy along a round trip is zero. This statement, trivial as it may seem, is another hallmark of potential energy. No other form of energy has this property.

7.14 As the ball moves upward, it slows down, reducing its kinetic energy. Because the Earth-ball system is closed, the gravitational potential energy of the system must therefore increase as the ball moves upward. The slowing down of the upward-moving ball means that the acceleration points down—the direction in which the gravitational potential energy of the Earth-ball system decreases. When the ball moves downward, it is again accelerated downward, its kinetic energy increases, and its potential energy decreases.

7.15 (*a*) It increases (see Checkpoint 7.14). (*b*) $\Delta U^G = (3.4\text{ kg})\cdot(9.8\text{ m/s}^2)(1.0\text{ m}) = 33\text{ J}$. (*c*) Your arm muscles burn up some chemical energy as you raise the book.

7.16 No. Even though you may hear people say that the quantity expressed by Eq. 7.21 is "the gravitational potential energy of the ball," it is incorrect to do so. It is not the ball by itself that possesses the potential energy. The gravitational potential energy is a property of the configuration of the Earth-ball system. Put another way, the ball by itself is not a closed system—it does not fall by itself. Consequently, the energy of the ball by itself is not constant.

7.17 With $e = 0$, Eq. 7.26 gives $\Delta E_{th} = \frac{1}{2}\mu v_{12,i}^2$. The reduced inertia is $\mu = (1000\text{ kg})(2000\text{ kg})/(1000\text{ kg} + 2000\text{ kg}) = 667\text{ kg}$, and the initial relative speed is $v_{12,i} = 50\text{ m/s}$. Thus $\Delta E_{th} = \frac{1}{2}(667\text{ kg})(50\text{ m/s})^2 = 8.3\times10^5\text{ J}$.

Chapter 8

8.1 Because the crate's velocity is constant, its momentum is also constant. So the change in its momentum is zero, and therefore the rate of change in its momentum is also zero. Surprised? Read on.

8.2 (*a*) As you set the crate in motion, the acceleration and the change in momentum point in the direction of travel. Because the vector sum of the forces exerted on the crate is equal to the time rate of change in its momentum, the vector sum points in the direction of travel. This makes sense: To make the crate move faster in a certain direction, you must push it in that direction. (*b*) When the crate is slowing down, all the signs are reversed: The acceleration points opposite the direction of travel, and so the change in momentum and the vector sum of the forces point opposite the direction of travel. Again, this makes sense because in order to stop the crate, you have to exert a force in the direction opposite its direction of travel. (*c*) When the crate is at rest—whether you are pushing against it or not—the vector sum of the forces exerted on it is zero, and so there is no direction involved.

8.3 (*a*) Because both collisions have the same initial and final velocities, it suffices to look at just one. To see whether the momentum of the two-cart system is constant, we calculate the change in momentum of each cart:

$$\Delta v_{1x} = v_{1x,f} - v_{1x,i} = (+0.80\text{ m/s}) - 0 = +0.80\text{ m/s}$$

$$\Delta v_{2x} = v_{2x,f} - v_{2x,i} = (+0.20\text{ m/s}) - (+0.60\text{ m/s}) = -0.40\text{ m/s}$$

$$\Delta p_{1x} = m_1\Delta v_{1x} = (0.12\text{ kg})(+0.80\text{ m/s}) = +0.096\text{ kg}\cdot\text{m/s}$$

$$\Delta p_{2x} = m_2\Delta v_{2x} = (0.24\text{ kg})(-0.40\text{ m/s}) = -0.096\text{ kg}\cdot\text{m/s}.$$

The total change in momentum is therefore zero, $\Delta p_{1x} + \Delta p_{2x} = 0$, and so the momentum of the system does not change. (*b*) Whether or not the collision is elastic can be determined by looking at the relative speeds. Before the collision $v_{12i} = |v_{1x,i} - v_{2x,i}| = |0 - (+0.60\text{ m/s})| = 0.60\text{ m/s}$; after the collision $v_{12f} = |v_{1x,f} - v_{2x,f}| = |(+0.80\text{ m/s}) - (+0.20\text{ m/s})| = 0.60\text{ m/s}$. Therefore the relative speed is the same before and after the collision, $v_{12i} = v_{12f}$, and so $e \equiv v_{12f}/v_{12i} = 1$, and the collisions are indeed elastic. (You can obtain the same result by verifying that the kinetic energy of the two-cart system doesn't change.)

8.4 Yes. Because two colliding objects can be considered isolated, the momentum of the system remains unchanged regardless of the type of interaction. Therefore the changes in the momentum of the objects are equal in magnitude but opposite in direction, $\Delta\vec{p}_1 = -\Delta\vec{p}_2$, and so the objects exert equal forces on each other in opposite directions.

8.5 (*a*) The two magnitudes are equal. If that surprises you—how can something so small exert a force on something so large?—remember that any change in the book's momentum must be made up by an equal change of opposite sign in Earth's momentum, because the Earth-book system is isolated. (*b*) Yes. The fact that the book exerts a force $\vec{F}_{\text{by book on Earth}} = \Delta p_E/\Delta t$ means Earth's momentum must be changing. Its velocity must therefore be changing, and so it is accelerating. Because its inertia is so great, however, the acceleration is negligible: The Earth-book system is isolated and so $\Delta\vec{p}_E + \Delta\vec{p}_b = m_E\Delta\vec{v}_E + m_b\Delta\vec{v}_b = \vec{0}$, so that

$$\Delta\vec{v}_E = \frac{m_b}{m_E}\Delta\vec{v}_b \approx \vec{0},$$

which means that Earth's acceleration during this interaction is vanishingly small.

8.6 (*a*) Because the collision is totally inelastic, the final velocities are the same: $v_{bx,f} = v_{mx,f}$. Because the mosquito-bus system is isolated, the final momentum of the system must be the same as the initial momentum. The x component of the initial momentum is

$$\begin{aligned}p_{x,i} &= m_m v_{mx,i} + m_b v_{bx,i} = m_m(0) + m_b v_{bx,i}\\ &= 0 + (10^4\text{ kg})(+25\text{ m/s}) = +2.5\times10^5\text{ kg}\cdot\text{m/s}.\end{aligned}$$

The x component of the final momentum is

$$p_{x,f} = m_m v_{mx,f} + m_b v_{bx,f} = (m_m + m_b)v_{x,f},$$

where $v_{x,f}$ is the x component of the common final velocity. Substituting inertia values gives

$$p_{x,f} = (10^{-4}\text{ kg} + 10^{4}\text{ kg})v_{x,f} = p_{x,i} = +2.5 \times 10^{5}\text{ kg}\cdot\text{m/s}$$

or, solving for $v_{x,f}$,

$$v_{x,f} = +25\text{ m/s}.$$

It looks like the velocity of the bus hasn't changed at all. If you pretend that all the numerical values given are exact and keep more significant digits, however, you find that $v_{bx,f} = +24.99999975$ m/s, which means the bus velocity changes by a minuscule amount: $\Delta v_{bx} = -0.00000025$ m/s (no one in the bus notices!). The x component of the change in the mosquito's velocity is $\Delta v_{mx} = +25$ m/s.

(b) The x component of the change in the average acceleration of the bus is

$$a_{bx} = \Delta v_{bx}/\Delta t = (-0.00000025\text{ m/s})/(0.005\text{ s}) = -0.00005\text{ m/s}^2.$$

That of the mosquito is

$$a_{mx} = \Delta v_{mx}/\Delta t = (+25\text{ m/s})/(0.005\text{ s}) = +5000\text{ m/s}^2!$$

(c) The x component of the momentum changes are

$$\Delta p_{bx} = m_b \Delta v_{bx} = (10{,}000\text{ kg})(-0.00000025\text{ m/s}) = -0.0025\text{ kg}\cdot\text{m/s}$$

and

$$\Delta p_{mx} = m_m \Delta v_{mx} = (10^{-4}\text{ kg})(+25\text{ m/s}) = +0.0025\text{ kg}\cdot\text{m/s}.$$

Notice that these changes in momentum are equal in magnitude and opposite in direction, so that the momentum of the isolated system doesn't change.

(d), (e)

$$(F_{\text{by bus on mos}})_x = \Delta p_{mx}/\Delta t = (+0.0025\text{ kg}\cdot\text{m/s})/(0.005\text{ s}) = +0.5\text{ kg}\cdot\text{m/s}^2$$

$$(F_{\text{by mos on bus}})_x = \Delta p_{bx}/\Delta t = (-0.0025\text{ kg}\cdot\text{m/s})/(0.005\text{ s}) = -0.5\text{ kg}\cdot\text{m/s}^2.$$

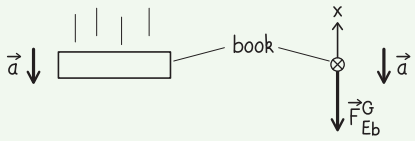
These two forces are equal in magnitude. Because of the large difference in inertia, however, the effect of this force on the bus is negligible, while on the mosquito the effect is enormous.

8.7 *Contact forces*: The only object in direct contact with the first magnet is the table, and so the table exerts a contact force on the magnet. There is no contact force with the other magnet because the two are not touching. Presumably the first magnet is also in contact with the surrounding air, and so there may also be a contact force exerted by the air on the magnet (in practice this force is negligible as long as the relative velocity of the magnet and the air is not too large). *Field forces*: Both gravitational and magnetic forces occur in this problem. There is a gravitational force exerted by Earth on the magnet and a magnetic force exerted by the second magnet. In principle, there is also a gravitational interaction between the two magnets and between the magnets and the table, but the forces due to these interactions are so small that we can safely ignore them.

8.8 No. Although the two forces are equal in magnitude and exerted in opposite directions, just as the two forces in an interaction pair are, they originate from different interactions and therefore are not part of the same interaction pair. The two forces in an interaction pair are exerted on different objects, whereas the two forces mentioned in this checkpoint are exerted on the same object (the book).

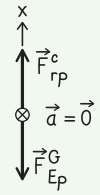
8.9 See Figure S8.9. During free fall, there is no contact force between the book and any other object. The only force exerted on the book is the gravitational force exerted by Earth (we ignore air resistance).

Figure S8.9



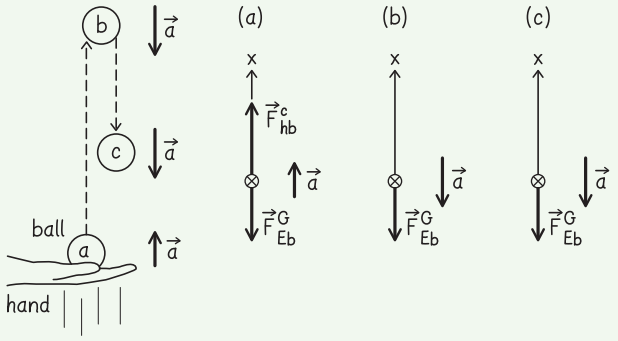
8.10 See Figure S8.10. The person is subject to one contact force: an upward force exerted by the ring (this is what keeps the person suspended). In addition, there is a downward gravitational force exerted by Earth.

Figure S8.10



8.11 (a) See Figure S8.11a. As you accelerate the ball upward, the force exerted by your hand on the ball is larger than the force of gravity exerted by Earth on the ball. (b) See Figure S8.11b. Once the ball is released, the only force exerted on it is that exerted by Earth (ignoring air resistance). (c) The forces exerted on the ball are the same on all parts of the trajectory. Therefore your drawing for part c should look exactly like your drawing for part b. Although the direction of the ball's velocity reverses during flight, the force exerted by Earth always points downward.

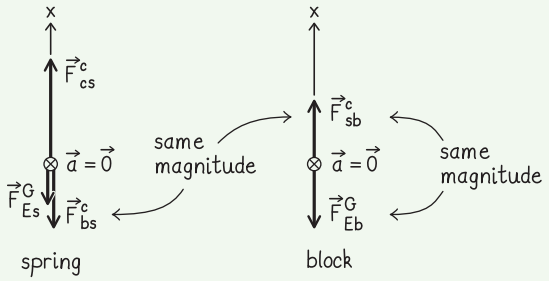
Figure S8.11



8.12 Figure S8.12 on the next page shows the free-body diagrams for the spring and the brick. The problem asks about the force exerted by the spring on the ceiling $\vec{F}^c_{sc}$, a force that forms an interaction pair with the force exerted by the ceiling on the spring. The brick is subject to a downward gravitational force exerted by Earth and an upward contact force exerted by the spring. Because the brick is at rest, we know that these two forces are equal in magnitude. The spring is subject to a downward gravitational force, a downward contact force exerted by the brick, and an upward contact force exerted by the ceiling. Because the spring is at rest we know that the forces must add up to zero and so the force exerted by the spring on the ceiling is equal to the combined gravitational forces exerted by

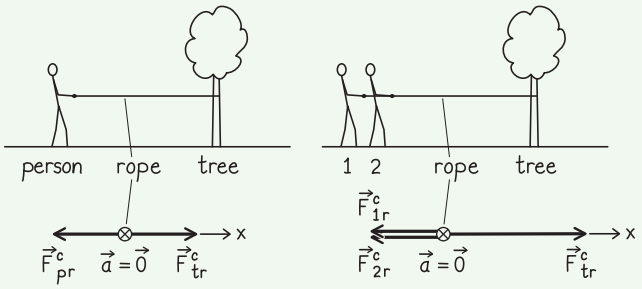
Earth on the spring and the brick. (This result makes sense because the ceiling has to support both of them.)

Figure S8.12



8.13 Figure S8.13 shows the free-body diagram for the rope. The tree now must exert a force of magnitude $2F$ to balance the forces exerted by the two people. This means a force of magnitude $2F$ is exerted on both ends of the rope, and so the tension is $\mathcal{T} = 2F$—twice as large as when the two people pulled on opposite ends.

Figure S8.13



8.14 (*a*) The force you exert on your friend gives her a constant acceleration given by Eq. 8.7, where m is the inertia of your friend and her skates. Estimating m to be 60 kg (adjust for your own estimate), you have $a = (200\text{ N})/(60\text{ kg}) = 3.3\text{ m/s}^2$. Her displacement in the 2.0s is $\Delta x = \frac{1}{2}a_x(\Delta t)^2 = \frac{1}{2}(+3.3\text{ m/s}^2)(2.0\text{ s})^2 = +6.7\text{ m}$. (*b*) When a person jumps off a wall, he is in free fall, and so the magnitude of his acceleration is 9.8 m/s^2. (*c*) When you put the person's inertia at 60 kg (adjust to your liking), the magnitude of the force exerted by Earth on the person is $F^G_{Ep} = (60\text{ kg})(9.8\text{ m/s}^2) = 5.9\times10^2\text{ N}$.

8.15 The two forces in an interaction pair are exerted on different objects and so never cancel when you consider the forces exerted on a single object. For example, if you push on a crate, the force you exert on the crate accelerates the crate. The other force in the interaction pair is a force exerted by the crate *on you*, not on the crate. The effect of this force exerted by the crate is to slow you down, not to cancel the force you exert on the crate.

8.16 (*a*) If you answered $m_E g$, think again! Does Earth have an upward acceleration g whenever you drop an object from a certain height? No; Earth doesn't budge. From the reciprocity of forces, however, you know that the gravitational force exerted by Earth on an object is equal in magnitude to the force exerted by the object on Earth. So the correct answer is $m_1 g$. (*b*) The effect this force has on Earth's acceleration is entirely negligible because of the enormous inertia of Earth (compare Checkpoint 8.6):

$$a_E = \frac{m_1}{m_E}g \approx 0.$$

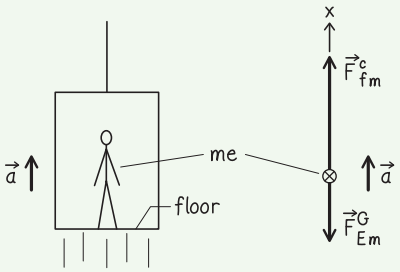
If you are wondering how it is possible that the force exerted by Earth on an object and the force exerted by the object on Earth both depend on the inertia of the object only, then you are very perceptive. I shall explain this when we discuss gravity in Chapter 13.

8.17 (*a*) See Figure S8.17. Note that the $\vec{F}^G_{Em}$ vector arrow is shorter than the $\vec{F}^c_{fm}$ arrow. You drew your diagram that way, right? (*b*) Because the elevator is accelerating upward, so are you. Therefore the vector sum of all the forces acting on you must point upward. This means that the contact force exerted by the floor of the elevator on you is larger than the downward gravitational force exerted on you. From Eq. 8.8,

$$\Sigma F_x = ma_x = F^c_{fmx} + F^G_{Emx}$$
$$F^c_{fmx} = ma_x - F^G_{Emx} = ma_x - (-mg) = m(a_x + g).$$

For an inertia $m = 60$ kg (substitute your own inertia), this yields a force of $(60\text{ kg})[(+1\text{ m/s}^2) + (9.8\text{ m/s}^2)] = 6.5\times10^2\text{ N}$.

Figure S8.17



8.18 (*a*) Stiffer. If k is large, a small displacement $x - x_0$ requires a large force (Eq. 8.18). A spring that compresses only a little under a large load is stiff. (*b*) Steel. Steel compresses less (small Δx) than foam rubber (larger Δx) under equal loads. So the spring constant for steel is larger than that for foam rubber.

8.19 (*a*) Smaller. Both objects have acceleration g, but the feather has much less inertia than the brick. The magnitude of the gravitational force is $F^G_{Eo} = mg$ (Eq. 8.17), and so the gravitational force exerted on the feather is smaller than that exerted on the brick. (*b*) The momentum is the same for both. The change in an object's momentum caused by a constant force—in other words, the impulse delivered to the object—is $\vec{F}\,\Delta t$. Because the same force is exerted on both objects for the same amount of time, $\Delta\vec{p}$ is the same for both. However the magnitude of $\Delta\vec{v}_{\text{feather}}$ is much larger than the magnitude of $\Delta\vec{v}_{\text{brick}}$ because of the difference in inertia.

8.20 (*a*) For a single object, the center-of-mass acceleration is equal to the particle's acceleration, $\vec{a}_{cm} = \vec{a}$, because the position of the center of mass coincides with the position x of the particle: $x_{cm} \equiv mx/m = x$ (Eq. 6.24). Finally for a single object there are no internal forces, so all forces are external and so $\Sigma\vec{F}_{ext} = \Sigma\vec{F}$. Consequently, Eqs. 8.46 and 8.47 are equivalent for a system made up of only one object. (*b*) Consider the motion of the system over a time interval $\Delta t = t_f - t_i$. If the center of mass initially moves at velocity $\vec{v}_{cm,i}$ at time t_i, then at the end of the time interval it has velocity

$$\vec{v}_{cm,f} = \vec{v}_{cm,i} + \vec{a}_{cm}\Delta t.$$

The change in the total momentum of the many-object system is thus

$$\Delta\vec{p} = \vec{p}_f - \vec{p}_i = m(\vec{v}_{cm,f} - \vec{v}_{cm,i})$$
$$= m(\vec{v}_{cm,i} + \vec{a}_{cm}\Delta t - \vec{v}_{cm,i}) = m\vec{a}_{cm}\Delta t.$$

After substituting Eq. 8.44 for the vector sum of the forces exerted on the system, you have your result:

$$\Delta\vec{p} = (\Sigma\vec{F}_{ext})\,\Delta t.$$

Chapter 9

9.1 (*a*) External, because the force is exerted from outside the system. (*b*) No, no, and no. Because the wall is hard, heavy, and anchored to the ground, it neither moves nor changes shape at a macroscopic scale. Because there is no friction involved, the wall's temperature does not change. (*c*) No. Its kinetic energy doesn't change because it doesn't accelerate; its potential energy doesn't change because it doesn't change shape or deform; its thermal energy doesn't change because it doesn't heat up. (*d*) No. The energy of the wall remains the same, and so there is no transfer of energy from you to the wall, which means you do no work on the wall no matter how hard you push.

9.2 Both forces do work on the ball because for both the point of application is at the ball, and this point moves as you launch the ball.

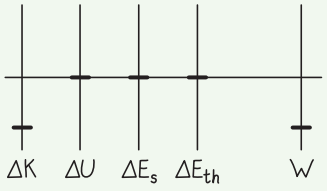
9.3 (*a*) As the ball moves upward, its kinetic energy decreases and its internal energy doesn't change. So the change in the ball's energy (the sum of the kinetic and internal energies) is negative. This means the work done on the ball is negative. (*b*) The ball's kinetic energy now increases, causing a positive change in energy, and so the work done on the ball is positive.

9.4 The directions of the force and the force displacement are consistent with the answers obtained by analyzing energy changes. In part *a* the only force exerted on the ball (the gravitational force) is downward, but the displacement is upward, and so the work done on the ball is negative. In part *b* both vectors are directed downward, and so the work done on the ball is positive.

9.5 (*a*) Yes. There are no changes in motion or state in the environment. (*b*) Positive. As the block compresses the spring, the force exerted by the block on the spring is directed toward the wall and the point of application of that force moves toward the wall. Force and force displacement pointing in the same direction means positive work. (Also, the energy of the system increases as the spring is compressed, and so the work done on it must be positive.) (*c*) Zero. The point of contact between wall and spring is not moving, and so the force displacement is zero, meaning that the work done by the wall on the system is also zero. (Indeed, the system comprising spring and block is closed, as we concluded in part *a*.)

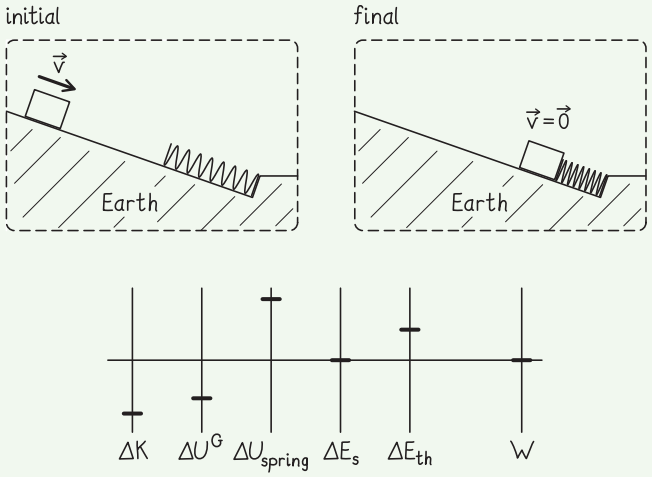
9.6 See Figure S9.6. The cart's kinetic energy decreases to zero, and there are no changes in the other forms of energy. The decrease in kinetic energy is the result of negative work done by the person on the cart. The force exerted by the person is directed to the left, and the force displacement is directed to the right, which is a second way of telling you that the work done by the person on the cart is negative.

Figure S9.6



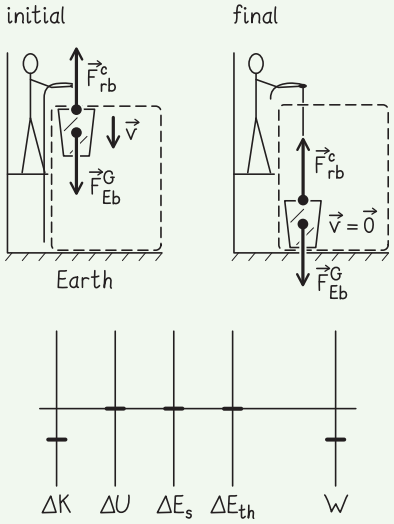
9.7 See Figure S9.7. As the block comes to rest against the compressed spring, the block's kinetic energy decreases, so you must draw the kinetic energy bar below the baseline. The system has two changes in potential energy: a change in gravitational potential energy ΔU^G due to the change in the block-Earth configuration and a change in elastic potential energy ΔU_{spring} due to the compression of the spring. You should draw two potential energy bars, one for each kind. The gravitational potential energy decreases, and the potential energy of the spring increases. The thermal energy increases because of friction. The lengths of the bars add up to zero because the system is closed, so $W = 0$.

Figure S9.7



9.8 See Figure S9.8. The only energy in the system that changes is the kinetic energy, which decreases to zero. The work done on the system must therefore be negative. This negative work is the sum of the work done by the rope on the system and the work done by Earth on the system. The work done by the rope on the system is negative (force upward, force displacement downward), and the work done by Earth on the system is positive (force and force displacement both downward). The work done on the system is negative because, in order for the system to slow down, the magnitude of the force exerted by the rope must be larger than that exerted by Earth.

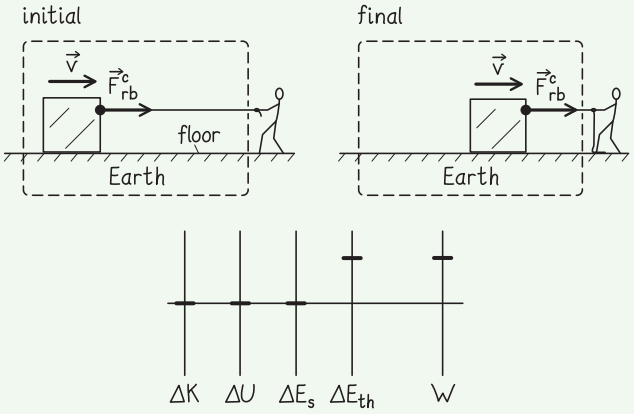
Figure S9.8



9.9 (*a*) See Figure S9.9 on the next page. The kinetic energy remains unchanged, but energy is dissipated because of the friction between box and floor. No changes occur in potential energy. The changes in thermal energy can be attributed to work done on the system. The person exerts external forces both on the box (via the rope) and on the floor (via his feet). Only the force displacement of the force exerted on the box is nonzero. Because the force displacement and the force point in the same direction, the work done by the person on the system is positive. (*b*) Because the system now includes the person/rope, the system is closed, which means no work is done on it and the W bar in the energy diagram is blank. As in the system chosen in part *a*, you have $\Delta K = 0$, $\Delta U = 0$, and $\Delta E_{\text{th}} > 0$ (friction

between box and floor increases E_{th} for the system). With the person/rope included as part of the system, the source energy bar must be filled in and it must be negative because muscle energy is used up as the person pulls.

Figure S9.9



9.10 The x component of the impulse delivered by the (constant) gravitational force is given by Eq. 8.25, $\Delta p_x = F^G_{\text{Eb}x}\Delta t = -m_b g \Delta t$. Because the acceleration is constant, $\Delta v_x/\Delta t = a_{x,\text{av}} = -g$, you can say that $\Delta t = -\Delta v_x/g$ and thus

$$\Delta p_x = -m_b g(-\Delta v_x/g) = m_b \Delta v_x = m_b(v_{x,f} - v_{x,i}).$$

9.11 Because Δx_F is larger than Δx_{cm} (see Figure 9.19), W in Eq. 9.16 is larger than ΔK_{cm} in Eq. 9.14. Because $\Delta E = W$, this means that ΔE is larger than ΔK_{cm}. This makes sense because only part of the work done on the system goes into changing K_{cm}. Another part goes into increasing the internal energy of the system. $W = \Delta E$ accounts for both the change in K_{cm} and the change in internal energy.

9.12 For a single particle, the position of the center of mass is the position of the particle (Eq. 6.24):

$$x_{cm} \equiv \frac{mx}{m} = x,$$

so the force displacement is equal to the displacement of the center of mass: $\Delta x_{cm} = \Delta x_F$. The center-of-mass translational kinetic energy is simply the particle's kinetic energy:

$$K_{cm} = \tfrac{1}{2}mv_{cm}^2 = \tfrac{1}{2}mv^2.$$

Therefore Eq. 9.14 becomes

$$\Delta K = (\Sigma F_x)\Delta x_F.$$

Because a particle has only kinetic energy, Eqs. 9.1 and 9.2 give $\Delta K = \Delta E = \Delta W$, so $\Delta W = (\Sigma F_x)\Delta x_F$, which is Eq. 9.9, the work equation for a single particle.

Now for reducing Eq. 9.18 to Eq. 9.9. Equation 9.18 is for a many-particle system. Because you want to use this equation for a single-particle system, there is only one particle experiencing the (external) force and so only one force displacement. That means you can take the factor Δx_{Fn} out of the summation in Eq. 9.18 and drop the subscript n. The equation is then identical to Eq. 9.9.

9.13 If the force is constant—in other words, if $F_x(x) = F_x$—it can be pulled out of the integral:

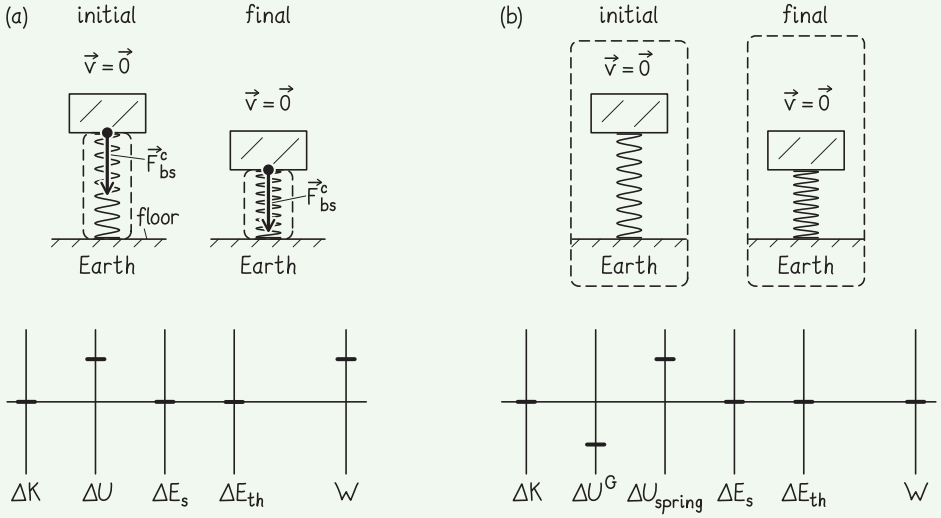
$$W = \int_{x_i}^{x_f} F_x(x)\,dx = F_x\int_{x_i}^{x_f} dx = F_x(x_f - x_i) = F_x\Delta x,$$

which is identical to Eq. 9.8.

9.14 See Figure S9.14. (*a*) This system is not closed, and the brick exerts a force on the system that has a nonzero force displacement. As it moves downward, the brick does positive work on the spring. This work ends up as elastic potential energy of the spring. (The force exerted by the floor on the spring does no work on the spring.) (*b*) The system is closed, so no work is done on it. In addition, none of the four types of energy changes! The increase in the elastic potential energy of the spring is accompanied by a decrease in the gravitational potential energy of the brick. To show this, you should draw two separate potential energy bars.

9.15 mg is the magnitude of the constant force of gravity exerted by Earth *on the brick*, not a force exerted on the spring. The contact force between brick and spring is a variable force described by Hooke's law.

Figure S9.14



If you find it surprising that the brick doesn't push down on the spring with the same force as Earth pulls down on the brick, look at the free-body diagrams for the brick in Figure S9.15. At the instant it is released just above the uncompressed spring, the brick is subject to only the downward force of gravity; the spring exerts no upward force as long as it is not compressed (Figure S9.15*a*). Consequently the x component of the brick's initial acceleration is $a_x = -g$. As the brick moves downward, the upward force exerted by the spring on it increases according to Hooke's law (Eq. 8.20, $F^{c}_{sbx} = -k(x - x_o)$. In this problem you are dealing with the situation represented by the free-body diagram in Figure S9.15*b*. At the instant represented in Figure S9.15*b*, $\vec{F}^{c}_{sb}$ is nonzero but still smaller in magnitude than F^{G}_{Eb} (because the brick has not yet come to rest against the compressed spring):

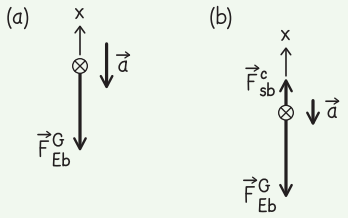
$$F^{c}_{bs} < F^{G}_{Eb} = mg.$$

Answering this checkpoint requires knowing $\vec{F}^{c}_{bs}$, which forms an interaction pair with $\vec{F}^{c}_{sb}$: $\vec{F}^{c}_{bs} = -\vec{F}^{c}_{sb}$, so

$$F^{c}_{bs} < F^{G}_{Eb} = mg.$$

The difference between the two force magnitudes F^{c}_{bs} and mg is what makes the brick accelerate downward. (If you wonder where I get these brain twisters, the answer is simple: As I wrote the example, I was wondering about this question myself.)

Figure S9.15



9.16 Although $\Delta x_{cm} = 0$ when the block comes to rest at its starting position, the distance it travels is nonzero: $d_{path} = 2(0.50\text{ m}) = 1.0\text{ m}$. All of the block's kinetic energy is converted to thermal energy, and so $\Delta E_{th} = K_i = \frac{1}{2}mv^2 = \frac{1}{2}(0.50\text{ kg})(1.0\text{ m/s})^2 = 0.25\text{ J}$. From Eq. 9.28, you get $F^{f}_{sb} = \Delta E_{th}/d_{path} = (0.25\text{ J})/(1.0\text{ m}) = 0.25\text{ N}$.

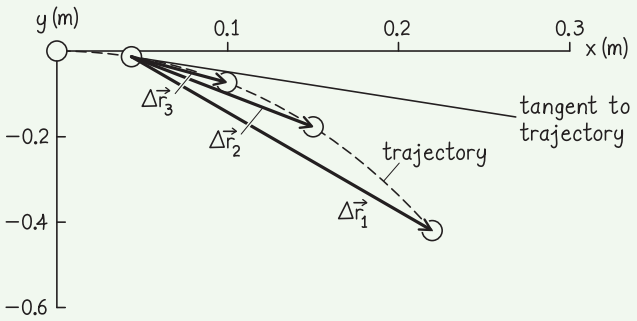
9.17 (*a*) Car: $(1.4 \times 10^8\text{ J})/(1800\text{ s}) = 78\text{ kW}$; plane: $(1.4 \times 10^8\text{ J})/(1\text{ s}) = 140\text{ MW}$. (*b*) Equation 7.19 tells you that, to raise the load, the gravitational potential energy of the load-Earth system must be increased by $\Delta U^G = (10\text{ kg})(10\text{ m/s}^2)(50\text{ m}) = 5000\text{ J}$. The athlete can deliver $500\text{ W} = 500\text{ J/s}$ and so takes $(5000\text{ J})/(500\text{ J/s}) = 10\text{ s}$. (*c*) $W_{\text{by engine}} = W_{\text{by athlete}}$ because the work done on the load does not depend on the rate at which it is done. In both cases, the work done on the load is equal to the change in the gravitational potential energy of the load-Earth system.

Chapter 10

10.1 (*a*) If your first reaction to this question was "In which reference frame?," you got the point! Before release, the ball is moving along with the cart, and so its initial velocity in the reference frame of the cart is zero. In the Earth reference frame, the ball's initial velocity is equal to that of the cart. (*b*) The speed is the magnitude of the velocity, which is the rate at which the displacement changes. Because of the cart's horizontal displacement, the ball's displacement from one frame of the film clip to the next is always greater in the Earth reference frame than in the cart's reference frame (see Figure 10.4). Therefore the ball's speed is higher in the Earth reference frame.

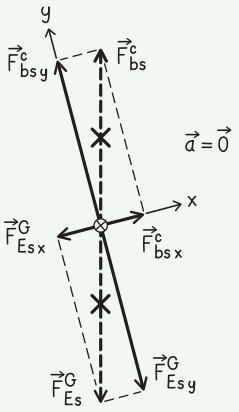
10.2 The ball's instantaneous velocity is the displacement over an infinitesimally small time interval. As Figure S10.2 shows, $\Delta\vec{r}$ approaches the tangent to the ball's trajectory as we make Δt smaller; $\vec{v}$ is thus tangent to the trajectory. Because of the downward curvature of the trajectory, $\vec{v}$ points above $\Delta\vec{r}$.

Figure S10.2



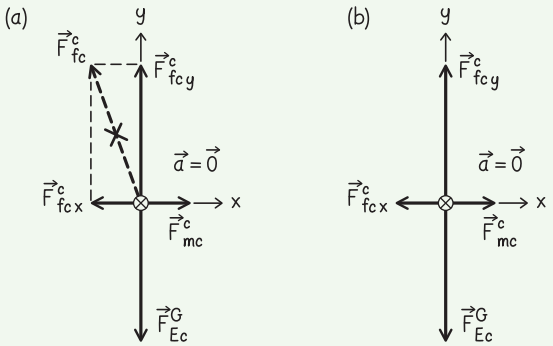
10.3 The suitcase is subject to a downward gravitational force and an upward contact force exerted by the belt. Because the suitcase's velocity is constant, its acceleration is zero, and so the vector sum of these forces must be zero. The two forces must therefore be equal in magnitude (Figure S10.3).

Figure S10.3



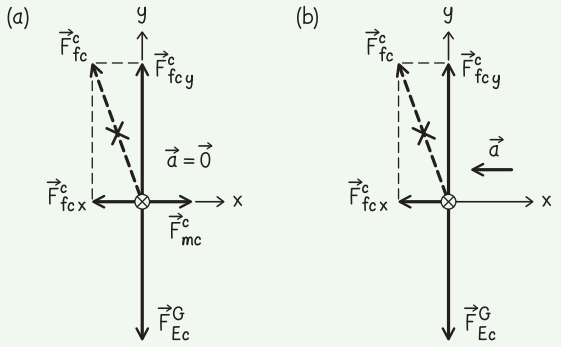
10.4 The cabinet participates in three interactions—with you, with the floor, and with Earth—and so is subject to three forces: a horizontal force due to your push $\vec{F}^{c}_{mc}$, a contact force $\vec{F}^{c}_{fc}$ exerted by the floor, and a downward force of gravity $\vec{F}^{G}_{Ec}$. Because the cabinet remains at rest, the vector sum of these forces must be zero. You know that your force is strictly horizontal and the gravitational force is strictly vertical. With these two forces being exerted on the cabinet, the only way for the vector sum of the forces exerted on the cabinet to be zero is for $\vec{F}^{c}_{fc}$ to have both a horizontal component $\vec{F}^{c}_{fcx}$ and a vertical component $\vec{F}^{c}_{fcy}$. The horizontal component of $\vec{F}^{c}_{fc}$ is therefore equal in magnitude to $\vec{F}^{c}_{mc}$, and the vertical component is equal in magnitude to $\vec{F}^{G}_{Ec}$. The free-body diagram for the file cabinet is shown in Figure S10.4*a*. (Alternatively, you can draw the two components of the contact force exerted by the floor as two separate forces, as shown in Figure S10.4*b*.)

Figure S10.4



10.5 (*a*) The cabinet is subject to the same three forces as in Checkpoint 10.4. Because it moves at constant speed, the vector sum of these forces is still zero, and so the free-body diagram is as shown in Figure S10.5*a*. (*b*) The instant you stop pushing, $\vec{F}^{\text{c}}_{\text{mc}}$ disappears but $\vec{F}^{\text{c}}_{\text{fc}\,y}$ and $\vec{F}^{G}_{\text{Ec}}$ must be as before because there is no vertical motion. You know that the cabinet comes to rest when you stop pushing, and so there must be a horizontal force that slows it *after* you stop pushing. This force is $\vec{F}^{\text{c}}_{\text{fc}\,x}$. The free-body diagram for the cabinet is shown in Figure S10.5*b*. Once the cabinet is at rest, $\vec{F}^{\text{c}}_{\text{fc}\,x}$ vanishes.

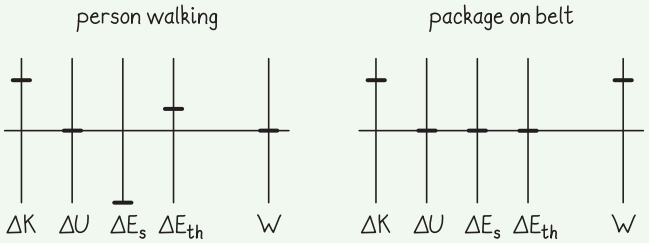
Figure S10.5



10.6 The motion of the two surfaces in contact relative to each other determines the type of friction: static friction between two surfaces not moving relative to each other and kinetic friction between two surfaces moving relative to each other. (*a*) Static, (*b*) kinetic, (*c*) kinetic, (*d*) kinetic, (*e*) static (as long as you are not slipping, your feet are at rest relative to the incline each time you put them on the ground while walking).

10.7 As the person accelerates, her kinetic energy increases. As you've just seen, this increase in kinetic energy is not due to work done on her by the force of static friction. The forces in the vertical direction add up to zero and so do no work either, so the work done on her is zero. What causes her to accelerate is that chemical energy in her muscles is converted to kinetic energy (and some of it is dissipated to thermal energy). So E_{th} increases and E_{s} decreases; the heights of the bars add up to zero (Figure S10.7). For the package, only its kinetic energy increases. Its shape and temperature remain the same, and so its internal energy does not change. The increase in kinetic energy is entirely due to work done by the conveyor belt on the package. The point of application of the force of friction exerted by the belt surface on the package is displaced in the direction of the force, so the work done on the package is positive, as Figure S10.7 shows.

Figure S10.7



10.8 Your final displacement $\Delta\vec{r}$ is the sum of the displacements $\Delta\vec{r}_1$ and $\Delta\vec{r}_2$ in Figure S10.8. You know the length and angular coordinate of $\Delta\vec{r}_1$ and the x and y components of $\Delta\vec{r}_2$. To add the two vectors, you must first determine the x and y components of $\Delta\vec{r}_1$:

$$\Delta x_1 = \Delta r \cos\theta = (1500\text{ m})(\cos 45°) = 1060\text{ m}$$

$$\Delta y_1 = \Delta r \sin\theta = (1500\text{ m})(\sin 45°) = 1060\text{ m}.$$

Adding the components of $\Delta\vec{r}_1$ and $\Delta\vec{r}_2$ gives

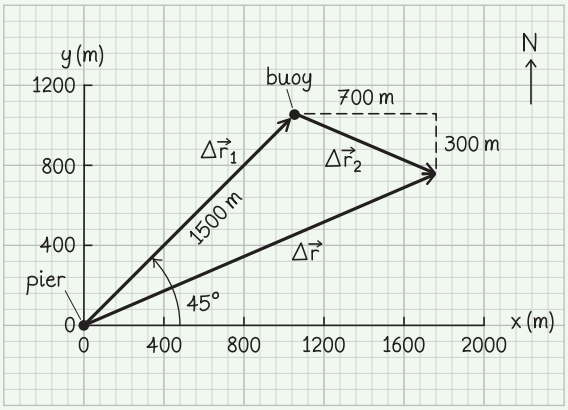
$$\Delta x = \Delta x_1 + \Delta x_2 = 1060\text{ m} + 700\text{ m} = 1760\text{ m}$$

$$\Delta y = \Delta y_1 + \Delta y_2 = 1060\text{ m} - 300\text{ m} = 760\text{ m}.$$

The distance to the pier is thus

$$|\Delta\vec{r}| = \sqrt{\Delta x^2 + \Delta y^2} = 1.92 \times 10^3\text{ m}.$$

Figure S10.8



10.9 (*a*) According to Eqs. 10.3: $v_{x,\text{i}} = v_{\text{i}}\cos\theta$ and $v_{y,\text{i}} = v_{\text{i}}\sin\theta$. Substituting these values into the expressions we obtained in Examples 10.5 and 10.6 yields

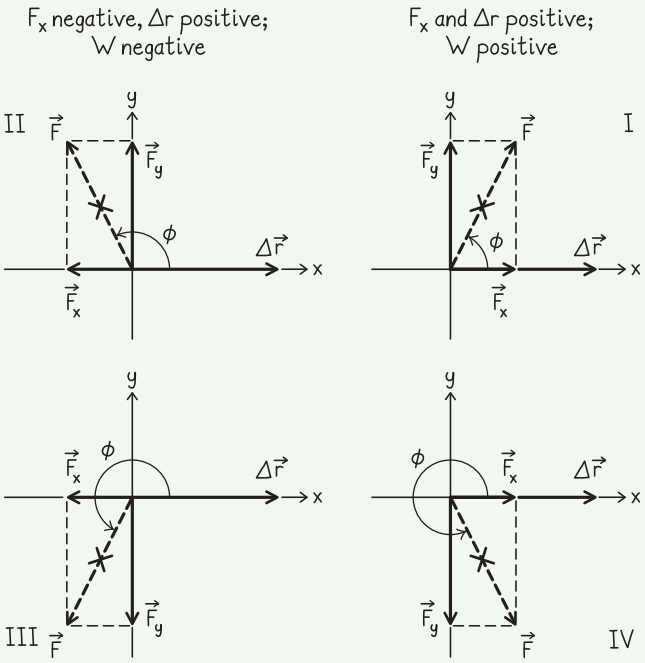
$$x_{\text{f}} = \frac{2v_{\text{i}}^2 \sin\theta\cos\theta}{g} \quad \text{and} \quad y_{\text{top}} = \tfrac{1}{2}\frac{v_{\text{i}}^2 \sin^2\theta}{g}.$$

(*b*) For a given launch speed v_{i}, y_{top} has its greatest value when $\sin^2\theta$ is maximal. This occurs when $\theta = 90°$, which makes sense because all of the projectile's speed goes into the upward motion when the projectile is launched straight up. (*c*) The greatest value of x_{f} is obtained when $\sin\theta\cos\theta = \frac{1}{2}\sin 2\theta$ is maximum. This occurs when $2\theta = 90°$ or $\theta = 45°$.

10.10 To answer this question, we need to compare the system's final and initial kinetic energies. Let the inertia of puck 1 be m; the inertia of puck 2 is then $2m$. The system's initial kinetic energy is $\frac{1}{2}m(1.8\text{ m/s})^2 + \frac{1}{2}\cdot 2m(0.2\text{ m/s})^2 = \frac{1}{2}m(3.3\text{ m}^2/\text{s}^2)$. The final kinetic energy is $\frac{1}{2}m(0.8\text{ m/s})^2 + \frac{1}{2}\cdot 2m(1.0\text{ m/s})^2 = \frac{1}{2}m(2.6\text{ m}^2/\text{s}^2)$. Because the system's final kinetic energy is not equal to the initial kinetic energy, the collision is not elastic.

10.11 To determine the sign of W, you need to know the sign of the scalar product for various relative positions of $\vec{F}$ and $\Delta\vec{r}$. To do so we choose a coordinate system that has its x axis aligned with $\Delta\vec{r}$ and then we place $\vec{F}$ in different positions (Figure S10.11). The scalar product is $F|\Delta\vec{r}|\cos\phi$, which can be written as $|\Delta\vec{r}|F\cos\phi = |\Delta\vec{r}|F_x$ and which is the product of the magnitude of $\Delta\vec{r}$ and the x component of $\vec{F}$. Because $|\Delta\vec{r}|$ is always positive, the sign of the scalar product is determined by the algebraic sign of F_x. As you can see in Figure S10.11, F_x is positive when $\vec{F}$ is in quadrants I and IV and negative when $\vec{F}$ is in quadrants II and III. So W is positive when $0 < \phi < 90°$ or $270° < \phi < 360°$. For any other angle, F_x and W are negative. According to Section 9.2, W is positive when $\vec{F}$ and $\Delta\vec{r}$ point in the same direction and negative when they point in opposite directions. So Eq. 10.40 is in complete agreement with our earlier definition of when work is positive and when it is negative.

Figure S10.11



10.12 (*a*) That the brick slides more easily on ice tells you that the force of friction is smaller in the brick-ice case than in the brick-wood case, and so μ_s is smaller for the brick on ice. (*b*) Substituting $\mu_s = 1$ into Eq. 10.46 gives $(F^s_{sr})_{max} = F^c_{sr}$ for the maximum force of static friction exerted by the surface on the runner. This is the maximum horizontal force the surface can exert on the runner. Because the runner is not accelerating up or down, the magnitude of the normal force must be equal to the magnitude of the force of gravity, and so $F^c_{sr} = F^G_{Er} = mg$, where m is the runner's inertia. Substituting this value into the equation of motion $\Sigma F_x = ma_x$ shows that the maximum acceleration is $a_x = +g$.

10.13 (*a*) If $\mu_s = 1$, then $\theta_{max} = 45°$ according to Eq. 10.53. (*b*) All objects start sliding well before this angle is reached, telling you that generally $\mu_s < 1$.

10.14 (*a*) The free-body diagram for the bricks is given in Figure 10.42. At the point where the bricks begin to slide, $F^G_{Eb\,x}$ is equal to the maximum value of F^s_{sb}, and so $F^G_{Eb\,x} + (F^s_{sb})_{max} = 0$. Because both $F^G_{Eb\,x}$ and $(F^s_{sb})_{max}$ are proportional to the inertia m of the bricks, the angle at which $F^G_{Eb\,x} + (F^s_{sb})_{max} = 0$ is independent of m. The single and double bricks therefore begin sliding at the same instant. (You can also see this directly from Eq. 10.53, which shows that θ_{max} is independent of m.) (*b*) Once the bricks begin to slide, kinetic friction replaces static friction. So the vector sum of the forces exerted on the bricks in the x direction is $\Sigma F_x = F^G_{Eb\,x} + F^k_{sb} = -mg \sin\theta + F^k_{sb}$. Using Eqs. 10.43 and 10.55, you can write this as $\Sigma F_x = -\mu_s F^c_{sb} + \mu_k F^c_{sb} = (\mu_k - \mu_s)F^c_{sb}$. The x component of the acceleration of the bricks is thus given by

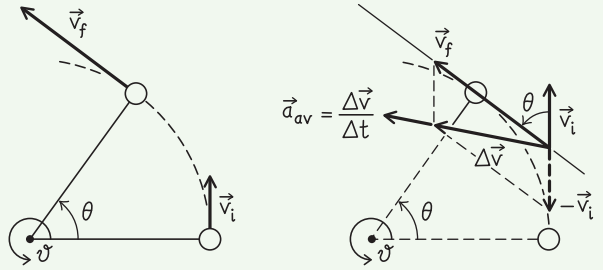
$$ma_x = \mu_k mg \cos\theta - mg \sin\theta$$

and is independent of inertia: $a_x = \mu_k g \cos\theta - g \sin\theta$. The single and double brick therefore both slide down with the same acceleration. These results can also be justified as follows: In parts *a* and *b* all forces exerted on the double brick are twice as large, but the double brick, having twice the inertia, requires twice as large a force to reach the same acceleration. Consequently there cannot be any difference between the accelerations.

Chapter 11

11.1 If $|\vec{v}_f| > |\vec{v}_i|$, the vector $\vec{v}_f$ has a greater magnitude than the vector $\vec{v}_i$. As a result, the change in velocity $\Delta\vec{v}$ and $\vec{a}_{av}$ now point not toward the center of the circle but more in the direction of motion, as shown in Figure S11.1 (compare with Figure 11.8*b*).

Figure S11.1



11.2 The centripetal acceleration is the same for both cubes, and so the force required to keep cube 1 going in a circle must be twice as large as the force required for cube 2. The maximum force of static friction, however, also scales with m—$(F^s_{tc})_{max} = \mu_s F^n_{tc} = \mu_s mg$—and so both cubes begin sliding at the same instant.

11.3 (*a*) Yes. In order to generate a vector sum of forces that points toward the center of the trajectory, the contact force exerted by the road surface on the bicycle (and that exerted by the bicycle on its rider) must be at an angle with the vertical. This can happen only when the bicycle leans into the curve. (*b*) No. The vertical component of $\vec{F}^c_{sb}$ is what prevents the bucket from falling to the ground: $F^c_{sb\,z} + F^G_{Eb} = 0$. If the rope is horizontal, then so is $\vec{F}^c_{sb}$ and consequently $F^c_{sb\,z} = 0$. Without this vertical component, the bucket cannot stay in the air.

11.4 (*a*) 3. The largest rotational inertia is obtained when the inertia of the pencil is as far as possible from the axis. For axis 3, half of the pencil is farther from the axis as for axis 2. (*b*) 1. For this axis all the inertia of the pencil is close to the axis.

11.5 (*a*) In 1.5 s the object covers an arc length of 4.5 m. Thus $\vartheta = (4.5\text{ m})/(2.0\text{ m}) = 2.3$. (*b*) The perimeter of the circle is $2\pi(2.0\text{ m})$. At 3.0 m/s, it therefore takes the object $2\pi(2.0\text{ m})/(3.0\text{ m/s}) = 4.2$ s to complete one revolution. (*c*) In 1.0 s the object moves through an arc of 3.0 m, and so the change in rotational coordinate in 1.0 s is $(3.0\text{ m})/(2.0\text{ m}) = 1.5$. The object begins at a polar angle of 90°, which corresponds to a rotational coordinate of $(90°)(2\pi\text{ rad}/360°)/(1\text{ rad}) = \pi/2$. So the final rotational coordinate is $\pi/2 + 1.5 = 3.1$.

11.6 The faster you go, the larger the inward force required to make you go around the curve. When the maximum force $(F^s_{sp})_{max}$ is reached, $\tan\theta_{max} = \mu_s$ (Eq. 11.53) and Eq. 1 in Example 11.4 becomes

$$\mu_s\, mg = m\frac{v^2}{r}$$

$$v = \sqrt{\mu_s gr} = 6.6\text{ m/s},$$

where I have used the value $\mu_s \approx 1$ for the coefficient of static friction between skates and road.

11.7 (*a*) The speeds of B and C are equal, but the circumference of C's trajectory is twice as large, and so it takes C twice as long to complete one revolution. Therefore the ratio of the rotational velocities is 2. (*b*) Because of the r^2 dependence of rotational inertia (see Eq. 11.30), C's rotational inertia is four times as large as B's. Therefore the ratio of the rotational inertias is $\frac{1}{4}$.

11.8 Yes. Using Eq. 11.31, you can write for the final rotational kinetic energy:

$$\tfrac{1}{2}I_f\omega_f^2 = \tfrac{1}{2}(I_f\omega_f)\omega_f = \tfrac{1}{2}(I_i\omega_i)\omega_f > \tfrac{1}{2}(I_i\omega_i)\omega_i = \tfrac{1}{2}I_i\omega_i^2.$$

Because his arms' centripetal acceleration must increase as he pulls them in, the force required to pull them in increases, requiring physical effort. Thus internal chemical energy in his body is converted to kinetic energy.

11.9 (*a*) Suppose the dumbbell has rotational velocity ω, with each puck moving at speed $v = r\omega = (l/2)\omega$. The kinetic energy of the dumbbell is equal to the sum of the kinetic energies of the two pucks:

$$K = 2\left(\tfrac{1}{2}mv^2\right) = mv^2.$$

Substituting the Eq. 11.11 value for v, you get

$$K = m\left(\tfrac{1}{2}l\omega\right)^2 = \tfrac{1}{4}ml^2\omega^2.$$

Knowing from Eq. 11.31 that $K = \frac{1}{2}I\omega^2$, you can say that

$$\tfrac{1}{4}ml^2\omega^2 = \tfrac{1}{2}I\omega^2.$$

Solve this expression for I and you see that, indeed, $I = \frac{1}{2}ml^2$. (*b*) No. The dumbbell rotates but has no momentum. Its axis is fixed to the table, so it cannot have any translational motion. Even though v is nonzero for each puck, $v_{\text{dumbbell}} = 0$ and therefore $p_{\text{dumbbell}} = 0$. Because the axis exerts a force on the dumbbell, the puck-dumbbell system is not isolated, and so it is not necessary for its momentum to be constant.

11.10 When the axis is at the end, you should choose the origin at that end, and so the integration boundaries in Example 11.9 become $x = 0$ and $x = l$. Therefore,

$$I = \frac{m}{l}\int_0^l x^2dx = \frac{m}{l}\left[\frac{x^3}{3}\right]_0^l = \tfrac{1}{3}ml^2.$$

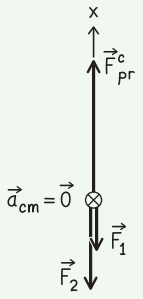
11.11 No, because drilling a hole removes material and its contribution to the rotational inertia, without changing the contribution of the remaining material: $I_{\text{solid}} = I_{\text{removed}} + I_{\text{remaining}} > I_{\text{remaining}}$. The m's in the two expressions for the rotational inertia are not the same: $m_{\text{hollow}} = \rho\pi l(R_{\text{outer}}^2 - R_{\text{inner}}^2)$, while $m_{\text{solid}} = \rho\pi lR_{\text{outer}}^2$. Substituting m_{hollow} into the expression for the rotational inertia of the hollow cylinder gives $I_{\text{hollow}} = \frac{1}{2}m_{\text{hollow}}(R_{\text{outer}}^2 + R_{\text{inner}}^2) = \frac{1}{2}\rho\pi l(R_{\text{outer}}^2 - R_{\text{inner}}^2)(R_{\text{outer}}^2 + R_{\text{inner}}^2) = \frac{1}{2}\rho\pi l(R_{\text{outer}}^4 - R_{\text{inner}}^4)$, which is *less* than $I_{\text{solid}} = \frac{1}{2}m_{\text{solid}}R_{\text{outer}}^2 = \frac{1}{2}\rho\pi lR_{\text{outer}}^4$.

11.12 About an axis through the center of mass. I is larger about any other axis because the term md^2 in Eq. 11.53 is always positive.

Chapter 12

12.1 (*a*) The rod is subject to four forces: the force of gravity $\vec{F}_{\text{Er}}^{G}$ (which we can ignore because the inertia of the rod is negligible compared to m_1 and m_2), the downward contact forces $\vec{F}_1$ and $\vec{F}_2$ exerted by the two objects, and an upward force $\vec{F}_{\text{pr}}^{\text{c}}$ exerted by the pivot that holds the rod in place. Because the rod remains at rest, the vector sum of these forces must be zero, and so the upward force must be equal in magnitude to the sum of the downward forces (Figure S12.1). (*b*) The rod is still subject to the same forces. The rod is pinned down at the pivot, so it cannot accelerate. Therefore $\vec{a}_{\text{cm}} = \vec{0}$ as before, and the free-body diagram is unchanged. (The rod rotates to one side, pulling object 1 up and lowering object 2, but this is not reflected in the free-body diagram.) (*c*) $r_1/r_2 = m_2/m_1$. (*d*) The rod rotates so that the end where object 1 was shoots up and the end holding object 2 falls down. (*e*) The rod rotates so that the end of the rod where the object of inertia $2m_1$ is suspended goes down and object 2 is pulled up. (*f*) The end of the rod where object 1 is suspended goes down. (*g*) The difference is the speed at which the rod rotates into a vertical position: fast in part *e* and slowly in part *f*.

Figure S12.1



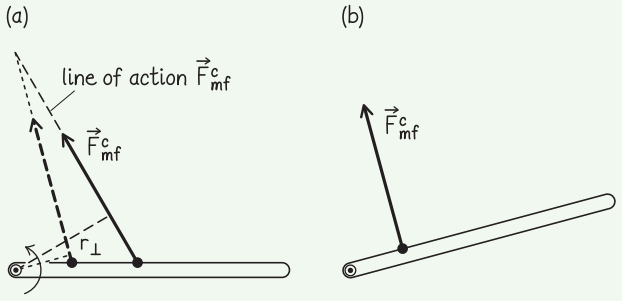
12.2 The seesaw remains at rest because the child causes a torque on the seesaw that is equal in magnitude to the torque you cause but tends to rotate the seesaw in the opposite direction.

12.3 (*a*) Because the torques caused by $\vec{F}_1$ and $\vec{F}_2$ cancel each other, the only way to prevent the rod from rotating is to make the torque caused by $\vec{F}_3$ zero. This can be achieved by aligning $\vec{F}_3$ along the long axis of the rod, either straight toward or straight away from the pivot. (*b*) The torque caused by $\vec{F}_1$ is $+r_{1\perp}F_1$; the torque caused by $\vec{F}_2$ is $-r_{2\perp}F_2 = -2r_{1\perp}F_2$. The lever arm distance of $\vec{F}_3$ is $r_{3\perp} = 3r_{1\perp}$, and so the torque caused by $\vec{F}_3$ is $+r_{3\perp}F_3 = +3r_{1\perp}F_1$ because $F_3 = F_1$. The sum of the torques is thus $+r_{1\perp}F_1 - 2r_{1\perp}F_2 + 3r_{1\perp}F_1$. Setting this sum to zero, I get $F_2 = 2F_1$. So, to balance the rod, you must increase the magnitude of $\vec{F}_2$ by a factor of 4.

12.4 If we neglect the very small effect of air resistance, the rotation of the wrench is steady. As you can verify by launching a stick or any other object, the object rotates in the same direction when it comes down. This shows that the rotational and translational motions are not coupled. The upward motion slows down because of the force of gravity, but this force does not affect the rotational motion because it does not cause a torque about the center of mass.

12.5 (*a*) The torque becomes greater because the lever arm distance of the force increases. See Figure S12.5*a*. (*b*) As the arm is raised, $\vec{F}_{\text{mf}}^{\text{c}}$ becomes more perpendicular to the forearm. See Figure S12.5*b*. This means that the torque caused by $\vec{F}_{\text{mf}}^{\text{c}}$ increases, and so the capacity to lift objects increases.

Figure S12.5



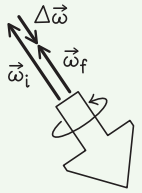
12.6 The acceleration of the center of mass is the same regardless of where the force is exerted. The acceleration of the center of mass of a system of particles is *always* given by $\sum\vec{F}_{\text{ext}} = m\vec{a}_{\text{cm}}$. See Section 8.12.

12.7 (*a*) Because the rotation slows down, the length of $\vec{\omega}$ decreases and so $\vec{\omega}_f$ is shorter than $\vec{\omega}_i$. The vector $\Delta\vec{\omega} \equiv \vec{\omega}_f - \vec{\omega}_i$ points from the tip of $\vec{\omega}_i$ to the tip of $\vec{\omega}_f$, which means the direction of $\Delta\vec{\omega}$ is opposite the direction of $\vec{\omega}$ (Figure S12.7). (*b*) Yes, because rotational acceleration is defined as (Eq. 11.12)

$$\alpha_\vartheta \equiv \lim_{\Delta t\to 0}\frac{\Delta\omega_\vartheta}{\Delta t}.$$

Because $1/\Delta t$ is a scalar, $\vec{\alpha}$ points in the same direction as $\Delta\vec{\omega}$.

Figure S12.7



12.8 The fact that the forearm is not moving tells you that $\Sigma\tau_{\text{ext}} = 0$. This yields

$$\frac{\ell}{5}F^{\text{c}}_{\text{mf}y} - \frac{\ell}{2}F^{G}_{\text{Ef}} - \ell F^{\text{c}}_{\text{bf}} = 0.$$

Dividing both sides by $\ell/5$ and solving for $F^{\text{c}}_{\text{mf}y}$ yields

$$F^{\text{c}}_{\text{mf}y} = \tfrac{5}{2}F^{G}_{\text{Ef}} + 5\,F^{\text{c}}_{\text{bf}}. \quad (1)$$

The condition that the vector sum of the forces exerted in the vertical direction must be zero yields

$$\Sigma F_y = F^{\text{c}}_{\text{mf}y} + F^{\text{c}}_{\text{hf}y} - F^{G}_{\text{Ef}} - F^{\text{c}}_{\text{bf}} = 0. \quad (2)$$

Substituting Eq. 1 into Eq. 2 and solving for $F^{\text{c}}_{\text{hf}y}$ yields

$$|F^{\text{c}}_{\text{hf}y}| = |-\tfrac{3}{2}F^{G}_{\text{Ef}} - 4F^{\text{c}}_{\text{bf}}| = \tfrac{3}{2}F^{G}_{\text{Ef}} + 4F^{\text{c}}_{\text{bf}}.$$

Note that the magnitude of the force that needs to be supplied by the biceps muscle is *much* greater than the sum of F^{G}_{Ef} and F^{c}_{bf}: $F^{\text{c}}_{\text{mf}} > |F^{\text{c}}_{\text{mf}y}| = \frac{5}{2}F^{G}_{\text{Ef}} + 5F^{\text{c}}_{\text{bf}} \gg F^{G}_{\text{Ef}} + F^{\text{c}}_{\text{bf}}$. This is the price one pays for the versatility and mobility of the forearm: If a load were suspended directly from the biceps muscle, our lifting capacity would increase by at least a factor of 5! Note furthermore that the downward force exerted by the humerus on the forearm is also very great. This great downward force is necessary to counter the upward force exerted by the biceps muscle on the forearm. Without this downward force, the muscle would simply pull the forearm up rather than rotate it.

12.9 (*a*) No. A nonzero vector sum of forces would cause the disc's center of mass to accelerate, and we know that the disc remains in the same place, so $\vec{a}_{\text{cm}} = \vec{0}$. (*b*) No. Even though the *vector sum* of the forces exerted on the compact disc is zero, the individual forces cause a nonzero torque. The resulting nonzero torque is what gives the disc a rotational acceleration and increases its rotational kinetic energy.

12.10 (*a*) The cylindrical shell has greater rotational inertia than the solid cylinder (MR^2 versus $\frac{1}{2}MR^2$; see Table 11.3) and so, because $\tau_\vartheta = I\alpha_\vartheta$, it requires a greater torque to be accelerated rotationally. This torque is supplied by the force of static friction, so the frictional force exerted on the shell must be greater. (You can see this directly from Eq. 12.26: The shape factor c is greater for the cylindrical shell, and so F^{s}_{ro} is greater, too.) (*b*) The maximum value of the force of static friction is, from Eq. 10.46,

$$(F^{s}_{\text{ro}})_{\max} = \mu_s F^{n}_{\text{ro}} = \mu_s mg\cos\theta.$$

Substituting this result into Eq. 12.26 yields

$$m_s\, mg\cos\theta = \frac{mg\sin\theta}{1 + c^{-1}}.$$

Solving this expression for θ, you obtain

$$\tan\theta = \mu_s(1 + c^{-1}).$$

For the cylindrical shell, $c = 1$, and so with $\mu_s = 1$, $\tan\theta = 2$ and $\theta = 63°$. For the solid cylinder, $c = \frac{1}{2}$, and so with $\mu_s = 1$, $\tan\theta = 3$ and $\theta = 72°$. Beyond these angles the cylinders no longer can roll without slipping. Note that both angles are larger than the 45° angle above which a block would begin to slip down a ramp. (See part *a* of Checkpoint 10.13.)

12.11 (*a*) Yes, the force of static friction does cause a torque on the wheel. The direction of this torque is counterclockwise in Figure 12.40*a*, opposite the direction of the torque caused by the chain. (*b*) Because the wheel's rotational acceleration is clockwise, the sum of the torques must be clockwise. Thus the clockwise torque caused by the chain must be greater than the counterclockwise torque caused by the frictional force.

12.12 Yes. The component $\vec{F}_\perp$ does work on the cylinder because its point of application is displaced along the circular arc. The arc length displacement is $\Delta s_F = r\Delta\vartheta$, and so the work done on the cylinder is $F_\perp \Delta s_F = F_\perp r\Delta\vartheta = \tau\Delta\vartheta$. This is equal to the change in the rotational kinetic energy in Eq. 12.31. So, for a rigid object subject to a constant torque, the quantity $t\Delta\vartheta$ is equal to the work done on the object.

12.13 $\vec{v} = \vec{\omega}\times\vec{r}$. Because $\vec{\omega}$ and $\vec{r}$ are orthogonal, the magnitude of this product is ωr, in agreement with Eq. 11.11. You can use the right-hand rule in Figure 12.44 to verify that the vector product of $\vec{\omega}$ and $\vec{r}$ indeed gives the right direction for $\vec{v}$.

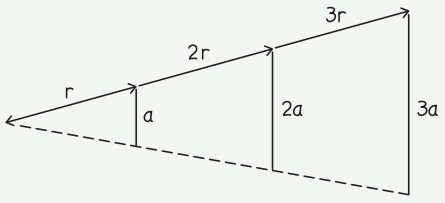
Chapter 13

13.1 (*a*) Rotational speed is the change in the rotational position divided by the time interval during which that change takes place, and so the object with the smaller period has the greater rotational speed. Because the Moon takes less time to complete one revolution, it has the greater rotational speed. Rotations you encounter in daily life—a spinning disc, a spinning wheel, the hands of a clock—rotate significantly faster and thus have significantly greater rotational speeds. (*b*) The speed is the distance traveled in 1 rev ($2\pi R$) divided by the period. Moon: $[2\pi(3.84\times10^8\text{ m})/(27.32\text{ days})]/(86{,}400\text{ s/day}) = 1.02\times10^3$ m/s. Earth: $[2\pi(1.50\times10^{11}\text{ m})/(365.26\text{ days})]/(86{,}400\text{ s/day}) = 2.99\times10^4$ m/s. Because of its much larger orbital radius, Earth's speed is considerably greater. (*c*) Moon: $a_c = v_{\text{M}}^2/R_{\text{M}} = (1.02\times10^3\text{ m/s})^2/(3.84\times10^8\text{ m}) = 0.00271\text{ m/s}^2$. Earth: $(2.98\times10^4\text{ m/s})^2/(1.50\times10^{11}\text{ m}) = 0.00595\text{ m/s}^2$. (*d*) These accelerations are much smaller than the acceleration due to gravity at Earth's surface: $g = 9.8\text{ m/s}^2$.

13.2 Earth's radius is about 6400 km, and so near the surface the distance to the center does not change much with altitude. For a plane 10 km above ground, the distance to the center of Earth is only $(10\text{ km})/(6400\text{ km}) = 0.016$ greater than that for an object on the ground. Because the gravitational force depends on the square of the distance, this increase results in a decrease in the strength of the force by a mere 2.4×10^{-6}!

13.3 (*a*) The light spreads out in a two-dimensional cone rather than a three-dimensional one, and the length of the straight line segment intersecting a two-dimensional wedge of light varies as $1/r$ (Figure S13.3), which means the force of gravity would exhibit a $1/r$ dependence. (*b*) The centripetal acceleration is still $a_c \propto R/T^2$. Because this acceleration is supplied by a gravitational force that is proportional to $1/r$, you have $R/T^2 \propto 1/R$ and so $T \propto R$.

Figure S13.3



13.4 For requirement (*i*) to hold, you must be able to interchange the indices in the expression for the gravitational force without changing the force magnitude: $F^{G}_{12} = F^{G}_{21}$. Requirement (*ii*) means that if you double one mass, the force exerted on the object must double. (*a*) No. This dependence is consistent with requirement (*i*) because $m_1 + m_2 = m_2 + m_1$, but not with requirement (*ii*) because $m_1 + 2m_2 \neq 2(m_1 + m_2)$. (*b*) Only m_1m_2 satisfies both requirements simultaneously. (*c*) If you take a very small piece m away, leaving $M - m$ behind, the attraction is proportional to $(M - m)m = Mm - m^2 \approx Mm$ (because the term m^2 is very small compared to Mm). So, the greater m is, the greater the force is. Symmetry thus suggests that the attraction is greatest when you divide the lump into two equal pieces. Mathematically you can prove this as follows: If you divide the lump into two pieces of masses cM and

$(M - cM)$, where $0 < c < 1$, then the gravitational force is proportional to $(M - cM)cM = M^2(c - c^2)$. To find the value for c at which the force is maximum, you must set the first derivative of this expression with respect to c equal to zero: $d[M^2(c - c^2)]/dc = M^2(1 - 2c) = 0$, so $1 - 2c = 0$ or $c = \frac{1}{2}$.

13.5 For the two forces to be equal in magnitude, you need

$$\frac{m_1 m_2}{r_{12}^2} = \frac{m_1 m_E}{R_E^2}.$$

You calculated the right side in Example 13.1, so

$$\frac{m_1 m_2}{r_{12}^2} = \frac{(70\text{ kg})^2}{r_{12}^2} = 1.0 \times 10^{13}\text{ kg}^2/\text{m}^2,$$

which yields

$$r_{12}^2 = \frac{(70\text{ kg})^2}{1.0 \times 10^{13}\text{ kg}^2/\text{m}^2} = 4.9 \times 10^{-10}\text{ m}^2,$$

or $r_{12} = 2.2 \times 10^{-5}$ m. This is the required separation between the *centers* of the bodies. Because the width of a human body is many orders of magnitude greater than this value, it is not possible to verify this prediction.

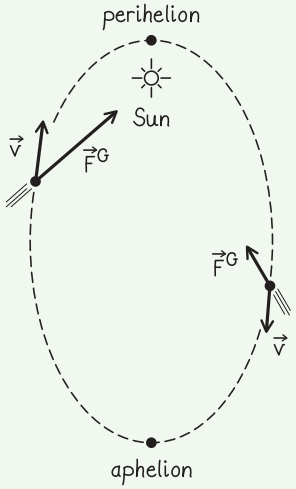
13.6 (*a*) To find the center of mass, use Eq. 6.24: $x_{cm} = (m_1 x_1 + m_2 x_2)/(m_1 + m_2)$. Let 1 be the Sun and 2 Earth. If you place the Sun at the origin, $x_1 = 0$ and x_2 equals the radius of Earth's orbit. Using the values in Table 13.1, you find

$$x_{cm} = \frac{0 + (5.97 \times 10^{24}\text{ kg})(1.50 \times 10^{11}\text{ m})}{(2.0 \times 10^{30}\text{ kg}) + (5.97 \times 10^{24}\text{ kg})}$$

$$= \frac{8.96 \times 10^{35}\text{ kg} \cdot \text{m}}{2.0 \times 10^{30}\text{ kg}} = 4.5 \times 10^{5}\text{ m}.$$

Compared with the radius of the Sun (7×10^8 m), this is a negligible distance: $x_{cm} = 0.00064 R_{Sun}$. (*b*) It speeds up because the force of gravity has a nonzero component parallel to the velocity, giving rise to an acceleration (Figure S13.6). The velocity keeps increasing until the comet reaches perihelion at the top of the ellipse. After passing that point, the comet slows down because the force of gravity causes an acceleration in the direction opposite the direction of the velocity. The slowing down continues until the comet reaches the point farthest from the Sun. At that point the velocity is a minimum.

Figure S13.6



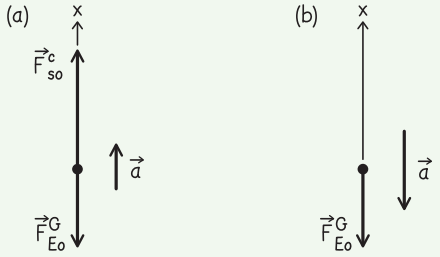
13.7 (*a*) The radius of Earth is about 6400 km, so the distance to Earth's center is increased by a factor of (6400 km + 300 km)/(6400 km) = 1.05. (*b*) Because the gravitational force changes as $1/r^2$, the force decreases by a factor of $1/(1.05)^2 = 0.91$. (*c*) The acceleration decreases by the same factor as the force: $0.91(9.8\text{ m/s}^2) = 8.9\text{ m/s}^2$. (*d*) The shuttle travels at such a high speed that the gravitational force exerted by Earth on it is just sufficient to provide its centripetal acceleration: $g = v^2/r$, with r the distance from the shuttle to the center of Earth.

13.8 (*a*) $F^G_{Eb} = mg = (1.0\text{ kg})(9.8\text{ m/s}^2) = 9.8$ N. (*b*) The scale exerts an upward contact force of magnitude F^c_{sb} on the brick. The brick is at rest, and so you know that the vector sum of the forces exerted on it is zero. Choosing the x axis in the upward direction, we have $\sum F_x = F^c_{sbx} + F^G_{Ebx} = F^c_{sb} + (-F^G_{Eb}) = 0$, and so $F^c_{sb} = F^G_{Eb}$. Furthermore, the contact force exerted by the brick on the scale and that exerted by the scale on the brick form an interaction pair, which means they are equal in magnitude, so $F^c_{bs} = F^c_{sb} = F^G_{Eb}$. (*c*) Two. Step 1: $F^c_{sbx} + F^G_{Ebx} = 0$; step 2: $F^c_{bs} = F^c_{sb}$. (*d*) Yes, because, as you found in part *b*: $F^c_{bs} = F^G_{Eb}$. (If you think this question is trivial, hold off judgment until after doing Checkpoints 13.9 and 13.10.)

13.9 (*a*) Inside the orbiting shuttle, $g = 8.9\text{ m/s}^2$ (see Checkpoint 13.7), and so $F^G_{Eb} = mg = (1.0\text{ kg})(8.9\text{ m/s}^2) = 8.9$ N—almost the same as at the surface of Earth (see Checkpoint 13.8). (*b*) Zero. The brick floats: no force is necessary to support it because everything around it has the same downward acceleration. (*c*) No. The scale measures nothing, even though Earth pulls on the brick nearly as hard as it does on Earth's surface. (*d*) The answer is different because the reference frame of the shuttle is accelerating toward the Earth reference frame.

13.10 (*a*) The same. Because the elevator is moving at constant velocity, its acceleration (and that of the spring scale and the object) is zero. Consequently the vector sum of the forces exerted on the object is zero, as it is on the ground, and so $\sum F_x = F^c_{sox} + F^G_{Eox} = 0$. (*b*) To accelerate upward, the vector sum of the forces exerted on the object must be nonzero and point upward (Figure S13.10*a*). If $a_x = +0.5g$, $\sum F_x = ma_x = +\frac{1}{2}mg$. Consequently $\sum F_x = F^c_{sox} + F^G_{Eox} = +\frac{1}{2}mg$, and so $F^c_{sox} = +\frac{1}{2}mg - F^G_{Eo} = +\frac{1}{2}mg - (-mg) = +\frac{3}{2}mg$. The scale reading is one and a half times what it is in a stationary elevator. (*c*) The object now has downward acceleration g and so is in free fall. The gravitational force exerted on it is thus the only force exerted on it (Figure S13.10*b*). The scale exerts no force on the object and so shows a zero reading.

Figure S13.10



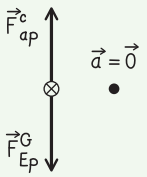
13.11 (*a*) In Checkpoint 13.7 you found that the shuttle's downward acceleration due to gravity is $g = 8.9\text{ m/s}^2$. The vertical distance Δh through which the shuttle falls in a time interval Δt is $\Delta h = \frac{1}{2}gt^2 = \frac{1}{2}(8.9\text{ m/s}^2)(1.0\text{ s})^2 = 4.5$ m. (*b*) If the shuttle remains 300 km above Earth, it is in a circular orbit, and so its centripetal acceleration is given by v^2/R, where v is its speed and R the radius of its orbit. Because the shuttle's acceleration is g, you have $v^2/R = g$, $v = \sqrt{gR} = \sqrt{(8.9\text{ m/s}^2)(6.7 \times 10^6\text{ m})} = \sqrt{5.7 \times 10^7\text{ m}^2/\text{s}^2} = 7.5 \times 10^3$ m/s.

13.12 (*a*) For the plane to be in free fall, it must have a downward acceleration of 9.8 m/s^2 (see Checkpoint 13.2). Starting with zero vertical velocity, a freely falling object drops in 40 s a distance $\Delta x = \frac{1}{2}gt^2 =$

$\frac{1}{2}(9.8\ \text{m/s}^2)(40\ \text{s})^2 = 7.8$ km. (*b*) The magnitude of the vertical speed is given by $v = gt = (9.8\ \text{m/s}^2) \times (40\ \text{s}) = 390$ m/s. (*c*) You can't tell. The instantaneous velocity of an object in free fall can be in any direction. Remember that a ball launched upward is in free fall while moving up as well as while moving down. At all times, its acceleration has the same constant downward value (see Section 3.3). The answers to parts *a* and *b* show that the plane *must* have been moving upward before going into free fall. If it had started horizontally, it would have lost almost 8 km of altitude and be dangerously close to the ground (airplanes normally fly at an altitude of about 10 km). Worse, its downward velocity at the end of the free fall would have been an amazing 1400 km/h, which is well above the maximum airspeed for most airplanes.

13.13 (*a*) The plane's speed is $v = (900\ \text{km/h})(1000\ \text{m/km})/(3600\ \text{s/h}) = 250$ m/s. Because the plane remains at a fixed altitude, it must follow a circular trajectory of radius $R \approx 6400$ km. Its centripetal acceleration is thus $v^2/R = (250\ \text{m/s})^2/(6.4 \times 10^6\ \text{m}) = 9.8 \times 10^{-3}\ \text{m/s}^2$. (*b*) See Figure S13.13. The centripetal acceleration of the plane is so small (about 0.1% of *g*) that you can safely ignore it, and so the free-body diagram is identical to what it would be if the plane were sitting on the ground. The plane is subject to two forces: a downward force of gravity and an upward force exerted by the air that is equal in magnitude to the downward force. (*c*) Because the airplane is not in free fall, its downward acceleration is only 0.1% of what it needs to be for the people to feel weightless.

Figure S13.13

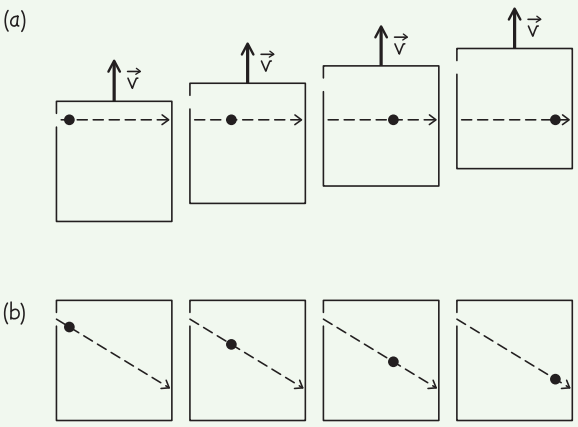


13.14 A passenger in a car that takes a right turn feels pushed outward from the turn, which means toward the left. To simulate this effect, the container should be tilted to the left.

13.15 (*a*) $\Delta t = \Delta x/v = (2.0\ \text{m})/(3.0 \times 10^8\ \text{m/s}) = 6.7 \times 10^{-9}$ s—just a few billionths of a second. (*b*) In the vertical direction you have $\Delta y = \frac{1}{2}a_y(\Delta t)^2$, $a_y = 2\Delta y/(\Delta t)^2$. To obtain a 1.0-mm deviation in the time interval obtained in part *a*, you need acceleration $a_y = 2(+1.0 \times 10^{-3}\ \text{m})/(6.7 \times 10^{-9}\ \text{s})^2 = +4.5 \times 10^{13}\ \text{m/s}^2$. This is such a phenomenally large acceleration that you cannot expect to observe this effect. (*c*) In 0.0010 s, a light beam travels $(3.0 \times 10^8\ \text{m/s}) \cdot (1.0 \times 10^{-3}\ \text{s}) = 3.0 \times 10^5$ m. In that time interval, it falls like an object in free fall, and so $\Delta y = \frac{1}{2}(-9.8\ \text{m/s}^2)(1.0 \times 10^{-3}\ \text{s})^2 = -4.9 \times 10^{-6}$ m—a tiny fraction of the diameter of a hair. Measuring this tiny displacement over a travel distance of 300 km is beyond current measurement accuracy.

13.16 (*a*) Zero. In the absence of both gravity and acceleration, the container and everything inside it are weightless. (*b*) The elevator moving at constant velocity has equal displacements in equal time intervals (Figure S13.16*a*). The path of the light pulse in the reference frame of the moving elevator is thus straight but at an angle (Figure S13.16*b*).

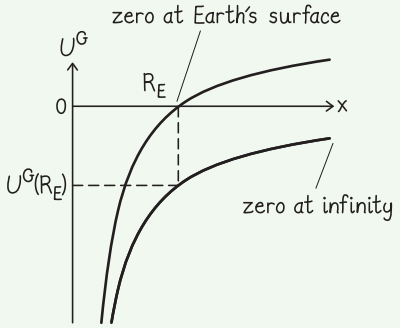
Figure S13.16



13.17 It increases because the gravitational force increases with decreasing distance r_{12}; see Eq. 13.1.

13.18 It yields values greater by an amount $|U(R_E)| = GmM_E/R_E$. See Figure S13.18.

Figure S13.18



13.19 (*a*) If we let the star be at rest at the origin, then the initial kinetic energies of the star and the object are zero. The initial gravitational potential energy U^G of the closed star-object system is given by Eq. 13.14 and is negative. The initial mechanical energy of the system $E_{mech} = K + U^G$ is thus negative. If we consider the star-object system to be closed, then its mechanical energy is constant, and so it is always negative because there is no dissipation of energy. (*b*) Because $\vec{F}^G_{so}$ is directed toward the star's center, the object accelerates in a straight line toward that center. The minimum kinetic energy is zero (see part *a*). Because the object accelerates as it moves toward the star, the maximum kinetic energy is achieved when the object reaches the star's surface. Conservation of energy tells you that $E_{mech,i} = E_{mech,f}$. From part *a*, $E_{mech,i} = U^G_i = -GMm/r$. At the star's surface, you have $E_{mech,f} = K_f + U^G_f = K_f - GMm/R_s$. So $K_f = GMm(r - R_s)/(R_s r)$. (Because r is the distance from the object to the center of the star, you have $r > R_s$, and so $K_f > 0$, as it should be.) (*c*) Yes. The angular momentum is $r_\perp mv$ (Eq. 13.36). Because the motion of the object is directed straight toward the center of the star, the lever arm distance about the center of the star remains zero throughout the object's motion. The angular momentum is thus both zero and constant. (*d*) The answer depends on the magnitude of the launch velocity: If $K_i < |U^G_i|$, then $E_{mech} < 0$; if $K_i = |U^G_i|$, then $E_{mech} = 0$; and if $K_i > |U^G_i|$, then $E_{mech} > 0$.

13.20 (*a*) At r_{max}, $U^G = E_{mech}$ and so $K = 0$. This means $v = 0$ and so $\vec{L} = \vec{0}$. (*b*) Conservation of angular momentum requires $\vec{L}$ to remain zero. This can happen only if $\vec{r} \times m\vec{v}$ is zero (or if v remains zero, which is impossible, because then nothing holds the object in place). This vector product is zero when $\vec{r}$ and $\vec{v}$ point in the same direction—in other words, when the object's motion is restricted to a straight line. (*c*) No. If $\vec{L} \neq 0$, $\vec{v}$ can never point in the same direction as $\vec{r}$ and neither of these can ever become zero. This means the object can never move all the way out to r_{max}. Although moving out that far does not violate conservation of energy, it does not allow the object to keep enough angular momentum.

13.21 (*a*) For the planet to move in a circular trajectory of radius R at constant speed v, the star must exert a gravitational force on the planet to cause the centripetal acceleration $a_c = v^2/R$. The gravitational force is thus $GMm/R^2 = ma_c$, or

$$\frac{GMm}{R^2} = m\frac{v^2}{R}.$$

To obtain the kinetic energy $\frac{1}{2}mv^2$, multiply both sides by $\frac{1}{2}R$:

$$\frac{1}{2}\frac{GMm}{R} = \frac{1}{2}mv^2.$$

(*b*) The mechanical energy in the circular orbit is the sum of the kinetic energy you just determined and the gravitational potential energy from Eq. 13.11:

$$E = \tfrac{1}{2}\frac{GMm}{R} - \frac{GMm}{R} = -\tfrac{1}{2}\frac{GMm}{R}.$$

Note that this result is the negative of the kinetic energy. So, for an object in a circular orbit, the kinetic energy is half the magnitude of the system's potential energy. (*c*) For a circular orbit, $e = 0$, and so from Example 13.5 you know

$$\frac{2E_{\text{mech}}L^2}{G^2M^2m^3} = -1,$$

which means $E_{\text{mech}} = -\tfrac{1}{2}G^2M^2m^3/L^2$. Substituting L^2 from Example 13.5 with $e = 0$ and $a = R$ gives $E_{\text{mech}} = \left(-\tfrac{1}{2}G^2M^2m^3\right)/(GMm^2R) = -\tfrac{1}{2}GMm/R$. (*d*) The two results are the same.

13.22 (*a*) At Earth's surface, the distance to the planet's center is equal to Earth's radius: $r_i = R_E$, and so Eq. 13.23 gives

$$\tfrac{1}{2}mv_{\text{esc}}^2 = \frac{Gmm_E}{R_E}$$

or $v_{\text{esc}} = \sqrt{2Gm_E/R_E}$. (*b*) $v_{\text{esc}} = [2(6.6738 \times 10^{-11}\ \text{N}\cdot\text{m}^2/\text{kg}^2)\cdot(5.97 \times 10^{24}\ \text{kg})/(6.378 \times 10^6\ \text{m})]^{1/2} = 1.12 \times 10^4\ \text{m/s}$. (*c*) No. As long as an object is fired above the horizon (so it doesn't hit the ground), only the magnitude of the velocity matters. Once the object has enough kinetic energy to reach infinity ($E = 0$), it does so regardless of the direction in which it is fired.

13.23 (*a*) The integral has two parts:

$$(r^2 - R^2)\int_{r-R}^{r+R}\frac{1}{s^2}ds + \int_{r-R}^{r+R}ds.$$

The first term yields

$$(r^2 - R^2)\left[-\frac{1}{s}\right]_{r-R}^{r+R} = (r^2 - R^2)\left[-\frac{1}{r+R} + \frac{1}{r-R}\right].$$

Bringing both terms under a common denominator yields

$$(r^2 - R^2)\frac{(R-r)+(r+R)}{r^2 - R^2} = 2R.$$

The second term of the integral yields

$$[s]_{r-R}^{r+R} = (r+R) - (r-R) = 2R.$$

Adding both parts together yields $4R$. (*b*) Equations 13.24–13.27 still hold when m lies inside the shell. From Figure S13.23, we see that

$$R\cos\theta = r + s\cos(\pi - \alpha) = r - s\cos\alpha,$$

which yields the same relationship between α and θ as in Eq. 13.28. Equation 13.29 is also still valid and so we obtain the same result as in Eq. 13.33. (*c*) The integration boundaries are now from $R - r$ (at $\theta = 0$) to $R + r$ (at $\theta = \pi$). The first part of the integral then yields

$$(r^2 - R^2)\left[-\frac{1}{s}\right]_{R-r}^{R+r} = (r^2 - R^2)\left[-\frac{1}{R+r} + \frac{1}{R-r}\right].$$

Bringing both terms again under a common denominator gives

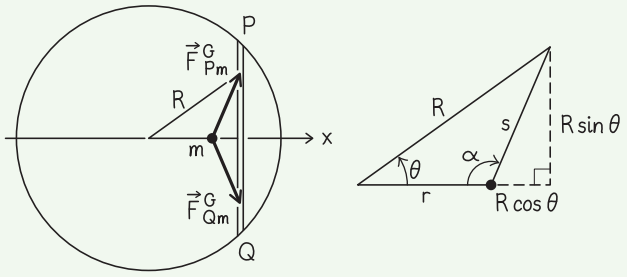
$$(r^2 - R^2)\frac{(r-R)+(R+r)}{R^2 - r^2} = -2r.$$

The second part of the integral yields

$$[s]_{R-r}^{R+r} = (R+r) - (R-r) = 2r.$$

Adding both parts together now yields zero. The gravitational attraction inside a uniform spherical shell is thus zero!

Figure S13.23



Chapter 14

14.1 (*a*) Because the light from event 1 reaches A before the light from event 2 reaches her, she observes event 1 first. (*b*) Because the light from event 1 and the light from event 2 reach B at the same instant, he observes the two events at the same instant.

14.2 (*a*) The light from each clock takes a finite time interval to travel from clock to observer. For clocks near the origin, this time interval is shorter than it is for clocks farther from the origin. This means (*ii*) is true: Nearby clocks display later time readings than distant clocks. (*b*) The clocks are synchronized and so by definition run at the same rate.

14.3 According to an observer moving along with the tube, the speed of the ball is l divided by Δt: $v_{\text{Mb}} = l/\Delta t$. Using the result $v_{\text{Mb}} = v_b + v_t$ from Example 14.3, you get $\Delta t = l/v_{\text{Mb}} = l/(v_b + v_t) = l/(2v)$.

14.4 (*a*) The observers agree that the signals do not reach the detector at the same instant because they both see that the blue light does not come on. (*b*) The observer at rest sees the two signals being emitted at the same instant. His reasoning for why the detector does not record simultaneity is that, because the detector is moving toward source 2 and away from source 1, signal 2 has a shorter distance to travel than signal 1. Therefore signal 2 reaches the detector first. The observer moving with the detector notes that the sources are equidistant from the detector but sees signal 2 arriving first. Because both signals travel at speed c_0 relative to him, he concludes that source 2 must have emitted its signal before source 1. The observers agree on the observation—no simultaneity—but have different explanations for it.

14.5 (*a*) Each signal takes $\Delta t = (5.0\ \text{m})/(3.00 \times 10^8\ \text{m/s}) = 1.7 \times 10^{-8}\ \text{s}$ to reach the detector. (*b*) $(20\ \text{m/s})(1.7 \times 10^{-8}\ \text{s}) = 3.3 \times 10^{-7}\ \text{m}$. (*c*) The travel time changes by $d/c_0 = (3.3 \times 10^{-7}\ \text{m})/(3.00 \times 10^8\ \text{m/s}) = 1.1 \times 10^{-15}\ \text{s} = 1.1\ \text{fs}$. (*d*) A 1.1-fs change in travel time is just barely measurable. (*e*) The detector moves $(1.0 \times 10^8\ \text{m/s})(1.7 \times 10^{-8}\ \text{s}) = 1.7\ \text{m}$. Light takes $(1.7\ \text{m})/(3.00 \times 10^8\ \text{m/s}) = 5.6 \times 10^{-9}\ \text{s}$ to travel this distance. Such a time interval can be measured, but note how fast the detector must be moving—$1.0 \times 10^8\ \text{m/s}$, about one-third the speed of light!—in order for the time interval to be long enough (if we can call $10^{-9}\ \text{s}$ "long") to be measurable.

14.6 (*a*) Greater than. Because the signal always travels at speed c_0, the longer distance the signal must travel according to B increases the duration of the cycle and therefore the period. (*b*) Increase. The faster

the clock moves relative to B, the farther the mirror moves away from its initial position above the location where the signal was emitted in one period of a clock at rest and so the period increases.

14.7 (*a*) Slow, because the mechanical watch takes longer to complete 1 minute. (*b*) Fast.

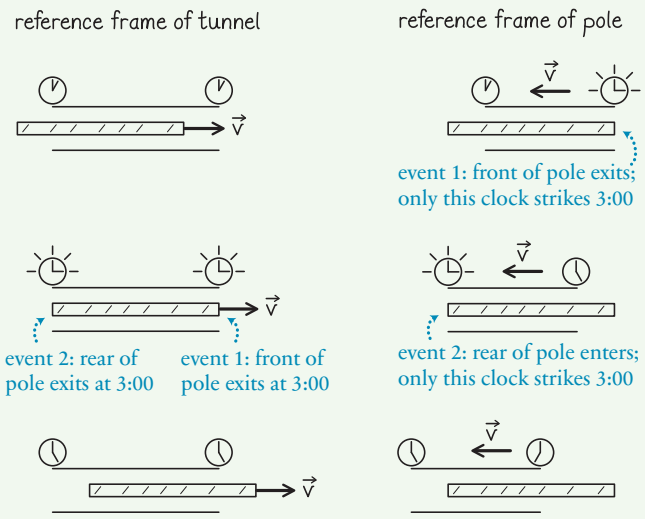
14.8 Yes and no. By moving at high speed, you could slow down your biological clock and therefore slow down the rate at which your body ages, but *only as measured by an observer who is not in your reference frame*. From your perspective, you see your biological clock running at its "regular" rate (that is, a rate that is the same as what you would measure if you remained at rest on Earth). To you, there is no slowing down of your biological clock, regardless of your motion relative to any observer. So, although you could lengthen your life in the eyes of an observer, it would be impossible to do so from your own point of view.

14.9 Figure S14.9*a* shows the situation viewed from B's reference frame. B starts accelerating relative to A when her clocks read 3 o'clock. Figure S14.9*b* shows the situation viewed from A's reference frame. Because the unit is moving, the rear clock is ahead of the front clock, and so A sees the rear clock accelerate before the front clock, which means A sees B's unit getting shorter.

14.10 No. See Figure S14.10. To the observer on the ground, the front end of the pole reaches the exit (event 1) at 3 o'clock and the rear end reaches the entrance (event 2) at the same instant. To the observer moving along with the pole, these two events do not occur simultaneously. To this observer, the tunnel is in motion, and so clocks at the two ends are desynchronized, with the clock at the exit ahead of the clock at the entrance. Consequently, this observer sees the front end of the pole exit the tunnel before the rear end enters and concludes that the pole is longer than the tunnel.

Is this disagreement a problem (like the problem of the clock fitting between the pegs in Figure 14.19)? Suppose the tunnel is equipped with doors that can close or open instantly. Initially the entrance is open and the exit is closed. When the front of the pole reaches the exit, the exit door opens. When the rear of the pole reaches the entrance, the entrance door closes. According to the observer on the ground, these two events occur at the same instant, and at that instant the pole fits exactly between the two doors. For the observer moving along with the pole, the front of the pole is at the exit before the rear is at the entrance, and so the exit door opens before the entrance door closes. The pole doesn't need to fit (and *can't* fit) in the tunnel.

Figure S14.10



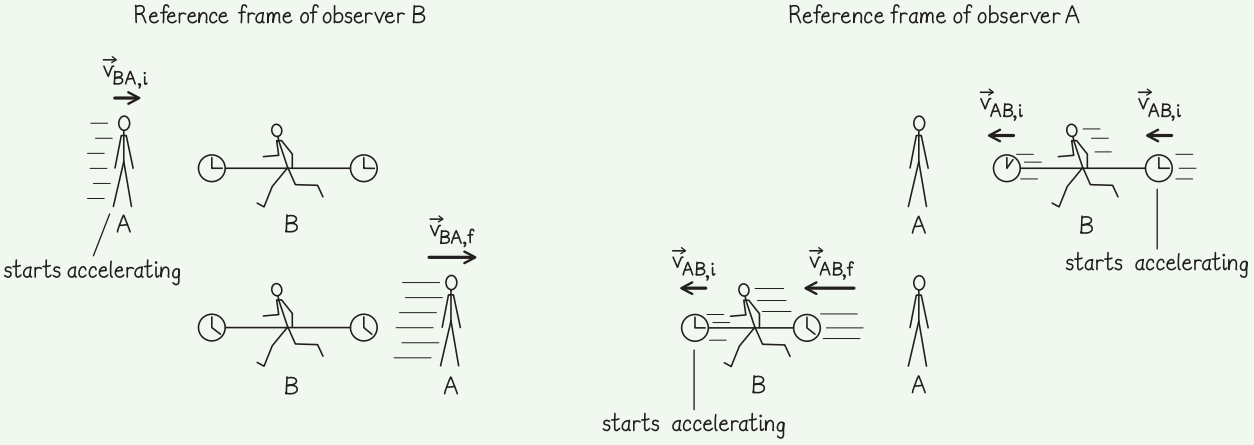
14.11 No, because the two observers measure different speeds for the object, and inertia depends on speed.

14.12 (*a*) When the spring is compressed, its internal energy increases, and so its mass increases. (*b*) When the coffee cools down, its internal energy decreases, and so its mass decreases.

14.13 (*a*) Because a particle has no internal energy, the mass is constant. Because the particle is accelerated, its kinetic energy increases. Therefore the energy of the system increases, and so the inertia increases. (*b*) As with any collision, we can consider the system of colliding objects to be closed, and so neither the energy nor the inertia can change. The internal energy must increase as the particle comes to rest in the target, and kinetic energy is converted to internal energy. Therefore the mass of the system increases.

14.14 If you use Eq. 14.6, your calculator will round the result to 1 for all but *d*. A better approach is to use Eq. 14.11 to calculate $\frac{1}{2}v^2/c_o^2$ and then add 1. (*a*) $\frac{1}{2}(30\text{ m/s})^2/(3.00\times10^8\text{ m/s})^2 = 5.0\times10^{-15}$, $\gamma = 1 + 5.0\times10^{-15}$. (*b*) $\gamma = 1 + (3.5\times10^{-13})$. (*c*) $\gamma = 1 + (5.6\times10^{-10})$. (*d*) For this part you cannot use Eq. 14.11 because the speed is not small. Equation 14.6 gives $\gamma = 1/\sqrt{1-(0.60)^2} = 1.25$.

Figure S14.9



14.15 From Eq. 14.17 you obtain $s^2 = [(3.00 \times 10^8 \text{ m/s}) \cdot (5.0 \times 10^{-7} \text{ s})]^2 - (120 \text{ m})^2 = 8.1 \times 10^3 \text{ m}^2$. This value equals $(c_o \Delta t_{proper})^2$. So $(\Delta t_{proper})^2 = (8.1 \times 10^3 \text{ m}^2)/(3.00 \times 10^8 \text{ m/s})^2$ and $\Delta t_{proper} = 3.0 \times 10^{-7}$ s.

14.16 According to an observer in the car, the 1-km length of road travels past the car at 100 km/h = 28 m/s. According to this observer, the measured length l_v is l_{proper}/γ (Eq. 14.28). The amount by which the length contracts is $l_{proper} - l_{proper}/\gamma = l_{proper}(1 - 1/\gamma) = l_{proper}(\gamma - 1)/\gamma$. For low velocities, γ is close to 1. According to Eq. 14.11, $\gamma - 1 \approx \frac{1}{2}v^2/c_o^2$, and so $(\gamma - 1)/\gamma \approx \frac{1}{2}v^2/c_o^2$. The difference in length is thus $l_{proper}(\frac{1}{2}v^2/c_o^2) = (1.0 \times 10^3 \text{ m})\frac{1}{2}[(28 \text{ m/s})/(3.00 \times 10^8 \text{ m/s})]^2 = 4.3 \times 10^{-12}$ m, which is a few orders of magnitude less than the smallest distance that can be measured. For all practical purposes, the road is still 1.0 km long.

14.17 Because of symmetry, the collision seen by A is identical to that seen by B. The time interval measured by A between the throw and catch of particle 1 is therefore equal to the time interval measured by B between the throw and catch of particle 2.

14.18 (*a*) Because $\vec{v}_2$ has a nonzero component along the x axis, A sees $v_2 > v_1$. (*b*) According to Eq. 14.40, $m_{v2} > m_{v1}$. Hence A concludes that inertia increases with speed. (*c*) From the perspective of B, $v_1 > v_2$. In going from reference frame A to reference frame B, the roles of particles 1 and 2 are interchanged, and so in reference frame B, Eq. 14.40 reads $m_{v1} = \gamma m_{v2}$. According to B, $m_{v1} > m_{v2}$, and so B, too, concludes that inertia increases with speed.

14.19 You know that for the electron at rest, m_v and m in Eq. 14.41 are identical, and so you want to know what speed v yields $\gamma = 1000$. Substituting this value into Eq. 14.9 yields:

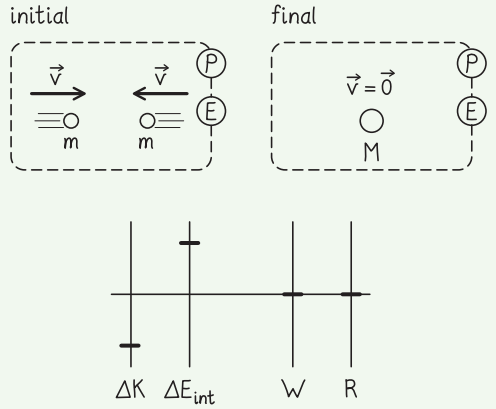
$$v^2/c_o^2 = 1 - 10^{-6}$$
$$v = \sqrt{1 - 10^{-6}}\, c_0 = 0.9999995 c_0.$$

14.20 (*a*) Because the final relative speed is zero, the collision is totally inelastic. (*b*) Using $\vec{p} = m\vec{v}$, you find $m_p v_{px,i} + m_e v_{ex,i} = 0$, $v_{e,i} = (m_p/m_e)v_{p,i} = (1836)(0.6c_0) = 1102c_0$!

14.21 $K = (\gamma - 1)mc_o^2 \approx (\frac{1}{2}v^2/c_o^2)mc_o^2 = \frac{1}{2}mv^2$.

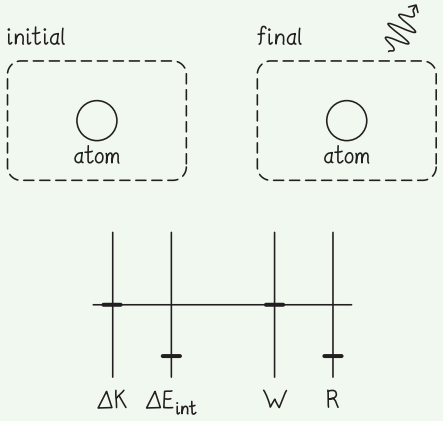
14.22 (*a*) See Figure S14.22. The collision converts kinetic energy to internal energy. (*b*) Kinetic energy is not constant because the final kinetic energy is zero. Mass is not constant because the increase in internal energy corresponds to an increase in mass. Energy is constant because the system is closed.

Figure S14.22



14.23 (*a*) See Figure S14.23. (*b*) The internal energy of the atom decreases because of the radiation emitted by it. The equivalence of mass and internal energy tells you that the mass of the atom decreases. (*c*) Because the atom is at rest, its inertia is equal to its mass. Because the mass decreases, the inertia also decreases.

Figure S14.23

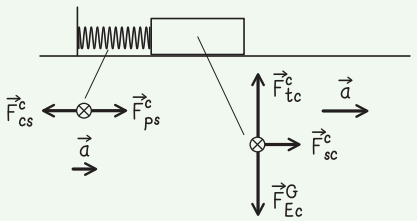


14.24 If you denote the initial energies as $E_{1i} = E_{2i} = E$, the energy law yields $E_{1i} + E_{2i} = 2E$. Because the momentum of the system is zero, we have from Eq. 14.57 $(2E)^2 = (300m_p c_o^2)^2$, or $E = 150m_p c_o^2$. From Eqs. 14.52 and 14.54, $K_{1i} = E - m_p c_o^2 = 149m_p c_o^2 = 2.2 \times 10^{-8}$ J, which is much less than the energy obtained in Example 14.13. (For this reason particle colliders use colliding beams rather than stationary targets.)

Chapter 15

15.1 (*a*) Two forces are exerted on the spring: one by the cart and the other by the post, and three forces are exerted on the cart: one by the spring, one by Earth, and one by the track. Figure S15.1 shows the free-body diagrams during compression (when the spring is stretched, the direction is reversed for the three horizontal forces in the diagrams). (*b*) None. The only forces that have a nonzero displacement along their line of action are the contact forces between the spring and the cart, but these are internal forces and so do no work on the system. (*c*) The contact force exerted by the cart on the spring points in the same direction as the force displacement, so the work done by this force on the spring is positive. This makes sense because, as the spring is compressed, its elastic potential energy increases. (*d*) The force exerted by the spring on the cart is directed in the direction opposite the force displacement, so the work done by this force on the cart is negative. Because $\vec{F}^c_{cs} = -\vec{F}^c_{sc}$ and the force displacement is the same for both forces, the work done by the spring on the cart is the negative of the work done by the cart on the spring.

Figure S15.1



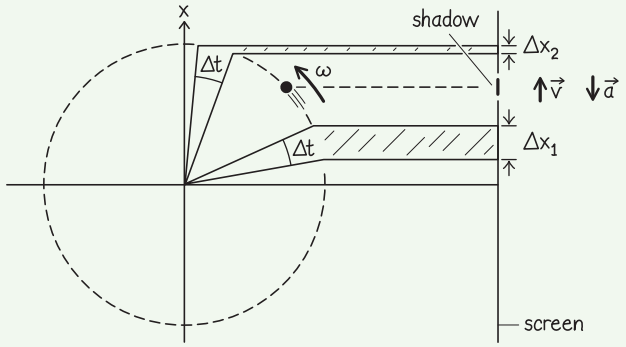
15.2 (*a*) Negative. Just before reaching the position where the spring is maximally stretched, the velocity is positive (points right), and just afterward, it is negative (points left). So the change in velocity and the acceleration are both negative. (*b*) It is greatest where the curvature of the $x(t)$ curve is greatest, when the spring is maximally stretched or compressed at the instants represented in Figure 15.2*a*, Figure 15.2*e*, and Figure 15.2*i*. It is smallest (zero) when the cart is at the equilibrium position at the instants represented in Figure 15.2*c* and 15.2*g*.

15.3 (*a*) *a*: tangential component of gravitational force, *b*: vertical component of elastic force in ruler, *c*: tangential component of gravitational force, *d*: vertical component of elastic force in string. (*b*) Gravitational potential energy in *a* and *c*, elastic potential energy in *b* and *d*.

15.4 (*a*) Factor of 4. From Chapter 9 you know that the potential energy of a spring is proportional to the square of the displacement from the equilibrium position. So, if the spring is compressed twice as much, the initial potential energy is four times as much. The initial kinetic energy is zero, so the energy is four times greater. (*b*) The initial compression determines the amplitude, and so the energy in the oscillator is proportional to the square of the amplitude.

15.5 (*a*) Upward, see Figure S15.5. (*b*) Downward. The two shaded regions in Figure S15.5 indicate that the shadow moves over increasingly smaller distances Δx during a given time interval Δt until it reaches the top. In other words, the shadow slows down. This means that the direction of the acceleration vector is opposite to that of the velocity vector.

Figure S15.5



15.6 (*a*) A pure sinusoidal function requires just a single term in the Fourier series, so the spectrum consists of a single peak at frequency $f = 1/T$. The peak height A is the square of the amplitude of the function. (*b*) As T increases, $f = 1/T$ decreases, and the single peak in the spectrum shifts to a lower frequency.

15.7 (*a*) At $x = x_0$ and for large x, $\Sigma F_x = 0$, and so these are equilibrium positions. Only the position at x_0 is a stable equilibrium position. (At very large x, the equilibrium is unstable because the object will tend to accelerate in the negative x direction.) (*b*) The shape of the curve tells you that the magnitude of the restoring force is greater for a negative displacement from x_0 than for a equal positive displacement.

15.8 If you did not set your calculator to work with radians before getting the five values, you might have concluded that the small-angle approximation $\sin\theta \approx \theta$ is not correct! Because radians measure ratios of arc lengths and radii (see Section 11.4), $\sin\theta \approx \theta$ applies only if θ is expressed in radians. Thus your first step is to convert each angle from degrees to radians, and then (with your calculator set at "rad") get the sine values. As Table S15.8 shows, the approximation is correct to better than 1% up to rotational positions corresponding to polar angles of 10°.

Table S15.8 Small-angle approximation

polar angle θ in degrees	polar angle θ in radians	$\sin\theta$	error (%)
1	0.0174533	0.0174524	0.0051
5	0.0872665	0.0871557	0.1270
10	0.1745329	0.1736482	0.5095
20	0.3490659	0.3420201	2.0600

15.9 (*a*) If the length of the pendulum is increased, the displacement of the pendulum bob for a given angle increases, but the restoring force remains the same. So the restoring force for a given displacement is smaller and thus the period is longer (for a mathematical proof see Example 15.6). (*b*) The smaller acceleration of gravity on the Moon decreases the restoring force exerted on the pendulum, and so the period is increased. (*c*) Neither the mass nor the spring constant is affected by the smaller gravitational attraction on the Moon, and so the object's period is the same on the Moon and on Earth.

15.10 (*a*) x positive, v_x positive, a_x negative. (*b*) x negative, v_x negative, a_x positive.

15.11 (*a*) Substituting the maximum displacement $x = A$ into Eq. 15.15, you get $U = \frac{1}{2}m\omega^2A^2$. (*b*) Yes. At maximum displacement, $K = 0$ and $E = U$.

15.12 (*a*) Yes. The velocity is given by the derivative of Eq. 15.6, as Eq. 15.7 shows. Because $\phi_i = \pi/2$, at $t = 0$, $v_x \propto \cos\pi/2 = 0$. (*b*) No. In Figure 15.26, for example, the cart moves to the left (negative x component of the velocity) even after crossing the position $x = 0$. Only after reaching $x = -A$ does the cart turn around and the x component of the velocity become positive.

15.13 (*a*) Because cart 2 remains in place after the first collision, the carts collide again when cart 1 returns to its initial position at $x = -15$ mm. Because the sine function is symmetrical about the maximum, this occurs after a time interval twice as long as that required to reach maximum compression. So $t = 2(0.17\text{ s}) = 0.34$ s. (*b*) Immediately after the second collision, cart 1 has zero velocity and then begins a new oscillation with an amplitude of 15 mm. Cart 2 moves away to the right at a constant speed of 0.10 m/s.

15.14 If the block is lifted above x_{eq}, Eq. 2 remains valid—the only difference now is that $x - x_{eq}$ is negative because x is on the other side of x_{eq}. The vector sum of the forces is then positive, reflecting the fact that the restoring force is now downward.

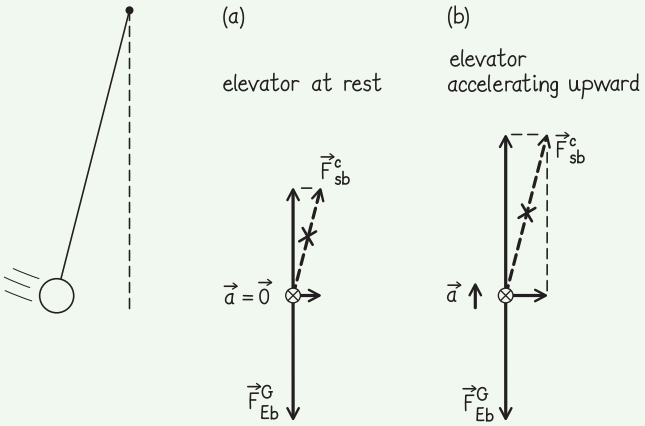
15.15 Decreasing the radius of the disk reduces its rotational inertia (which means it rotates more easily). Decreasing I increases ω (Eq. 15.29), and hence f increases (Eq. 15.4).

15.16 No. The oscillating rod experiment determines the acceleration due to gravity g, not G. The two are related (Eq. 13.4, $g = Gm_E/R_E^2$), but the relationship contains Earth's mass, which must be determined independently. [As you may remember, Earth's mass is determined from G, which is measured in the Cavendish experiment (see Section 13.5), not the other way around, see Example 13.3.]

15.17 Greater. The upward acceleration effectively increases g, which means that an object in your hand feels heavier. If g increases, the frequency of the pendulum increases too (see Example 15.6). An alternative way to see that f increases is to look at the free-body diagrams for the pendulum bob in an elevator at rest (Figure S15.17*a*, next page) and in an elevator that is accelerating upward (Figure S15.17*b*). For the elevator at rest (or moving at constant velocity), the upward

vertical component of the tensile force exerted by the string on the bob and the downward force of gravity add approximately to zero (ignoring the small vertical acceleration due to the pendulum motion). When the elevator accelerates upward, the tensile force exerted by the string on the bob must increase so that the vertical component of $\vec{F}^{c}_{sb}$ becomes greater than the gravitational force. This causes the pendulum bob to accelerate along with the elevator. If $\vec{F}^{c}_{sb}$ becomes greater, however, the horizontal component also increases, and so the restoring force for a given displacement increases, which increases f.

Figure S15.17

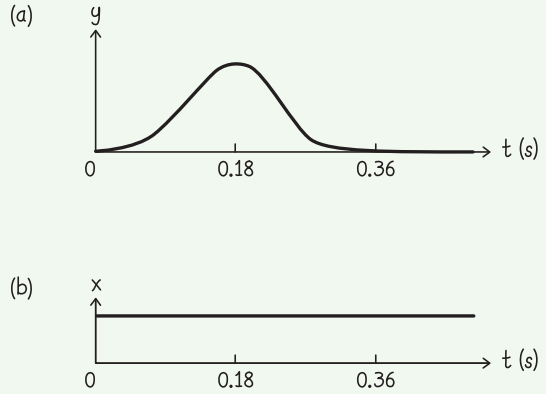


15.18 (*a*) The time-varying amplitude is given by $x_{max} = Ae^{-t/2\tau}$, where A is the initial amplitude. If the amplitude decreases by a factor of 3, $x_{max}/A = 1/3$, so $e^{-(4.0\text{ s})/2\tau} = 1/3$. Taking the natural logarithm (see Appendix B) of both sides, you get $-(4.0\text{ s})/2\tau = \ln(1/3) = -1.1$ and $\tau = (-2.0\text{ s})/(-1.1) = 1.8$ s. (*b*) $Q = \omega t = 2\pi f\,\tau = 2\pi(262\text{ Hz}) \cdot (1.8\text{ s}) = 3.0 \times 10^3$.

Chapter 16

16.1 (*a*) See Figure S16.1. (*b*) The x component of the bead's velocity is always zero because the bead moves in only the y direction. The y component of the bead's velocity is positive on the leading edge of the pulse and negative on the trailing edge.

Figure S16.1



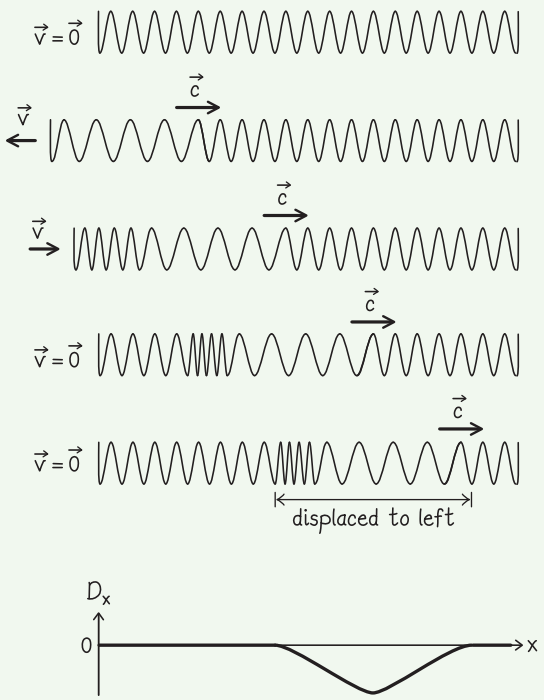
16.2 (*a*) Yes. Using a ruler, you can verify that the displacement of the pulse from one frame to the next is constant. (*b*) Between $t = 0$ and $t = 0.36$ s, the peak of the pulse moves over 16 beads, which corresponds to a displacement $\Delta x = (15)(5.0\text{ mm}) = 75$ mm, and so the wave speed is $c = (75\text{ mm})/(0.36\text{ s}) = 0.21$ m/s.

16.3 (*a*) 0.4 m. (*b*) 0.1 m. (*c*) 0. (*d*) Greater, because the pulse moves to the right and so the 1.5-m point on the string is about to be pulled upward.

16.4 (*a*) The point at $x = 0$ reaches its maximum height at $t = 0.50$ s; the point at $x = 1.0$ m reaches its maximum height at $t = 1.0$ s. So it takes the pulse 0.50 s to move 1.0 m. (*b*) $c = (1.0\text{ m})/(0.50\text{ s}) = 2.0$ m/s.

16.5 (*a*) Figure 16.16*b* plots the horizontal displacement of each coil on the spring from its *equilibrium* position. The coils between $x = 0.20$ m and $x = 0.40$ m are displaced from their equilibrium position and so $x = 0.30$ m is the midpoint of the pulse. The coil originally at $x = 0.30$ m is displaced 0.05 m to the right, as shown by the slanted line in Figure 16.6*a*. So, even though this coil appears at $x = 0.35$ m in Figure 16.6*a*, it really is at the midpoint of the pulse. (*b*) Yes, but the displacement is now negative. See Figure S16.5.

Figure S16.5

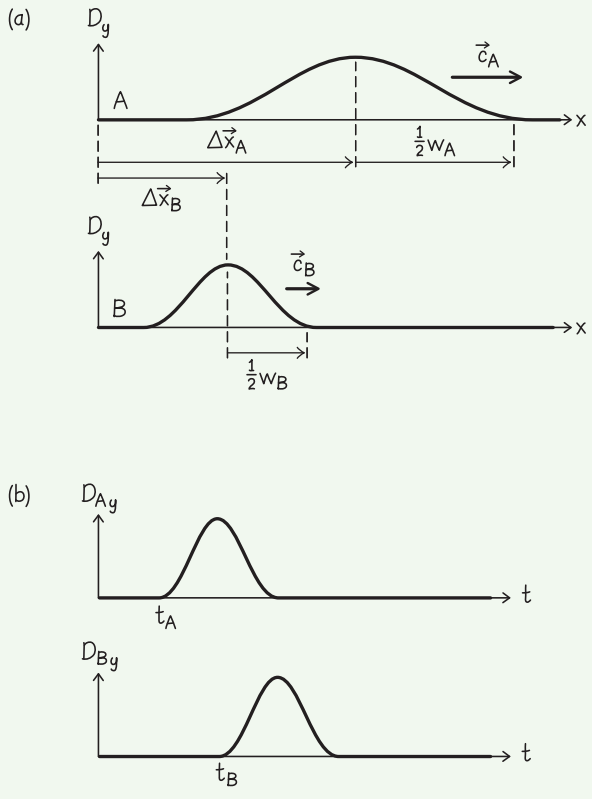


16.6 (*a*) At t_1, the free-body diagram for each bead to the right of bead 3 is identical to the one for bead 3. Because the accelerations of all these beads are zero, the forces exerted on each bead by its two neighboring beads are equal in magnitude and opposite in direction. In addition, you know that any two adjacent beads exert reciprocal forces on each other, and so at t_1 all the forces exerted on each other by the beads to the right of bead 3 must be equal in magnitude to F^{c}_{43} (5 N). To hold the right end of the string in place, some object must therefore be exerting a horizontal force of 5 N on that bead. (*b*) All beads move in only the vertical direction, and so their acceleration in the horizontal direction is always zero. You can thus use the same argument as in part *a* to show that the magnitudes of the horizontal components of all forces along the string are equal to 5 N. (*c*) The direction of the force exerted by the hand is tangent to

the direction of the string: up and to the left. (*d*) As you found in part *b*, the magnitude of the horizontal component is 5 N. Because the vertical component is nonzero, the magnitude of the force exerted by the hand must be greater than 5 N. (*e*) The horizontal component keeps the string taut; together with the force exerted by the wall at the right end, it provides the tension in the string. The vertical component accelerates the first bead vertically upward.

16.7 (*a*) The shape of the pulse does not change as it propagates along the string, and so each bead executes the same motion as the pulse passes through it. This means that each bead has velocity $\vec{v}_{\text{peak}}$ and acceleration $\vec{a}_{\text{peak}}$ when it is at its highest point, velocity $\vec{v}_{\text{half}}$ and acceleration $\vec{a}_{\text{half}}$ when it is halfway up to its highest point, and so on. Because the position of bead 4 at t_4 is the same as the position of bead 3 at t_3, $\vec{v}_4(t_4) = \vec{v}_3(t_3)$ and $\vec{a}_4(t_4) = \vec{a}_3(t_3)$. (*b*) They are the same. For each string, the motion of each point along the string is identical to the motion of the end you moved (provided the pulse doesn't change shape as it travels along the string). Because the end of A and the end of B move in identical fashion, all particles of *both* strings execute the same motion and so have equal velocities at equal displacements. (*c*) See Figure S16.7*a*. When the pulse on B has a displacement of $\Delta\vec{x}_{\text{B}}$, the displacement $\Delta\vec{x}_{\text{A}}$ of the pulse on A is twice as large because of the greater wave speed on string A. In addition—make sure you have not overlooked this—the pulse on A is twice as wide, for this reason: If, during the time interval it took your hand to move from initial position to maximum displacement, the pulse on A advanced a distance $\frac{1}{2}w_{\text{A}}$ (half the *width* of the pulse), then the pulse on B advanced only half as much: $\frac{1}{2}w_{\text{B}} = \frac{1}{2}\left(\frac{1}{2}w_{\text{A}}\right)$. (*d*) See Figure S16.7*b*. Each point carries out the same motion, taking the same amount of time to go up and then back down, and so the curves have the same shape. Because the pulse on B travels more slowly, the particle of B executes its motion at an instant t_{B} that is later than the instant t_{A} at which the particle of A executes its motion.

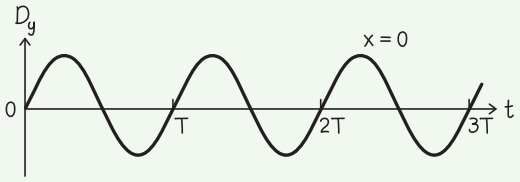
Figure S16.7



16.8 Because the tension in A is greater than that in B, the pulse travels faster on A than on B. This is the same as in Checkpoint 16.7, and so the sketches are the same. (*a*) See Figure S16.7*a*. (*b*) See Figure S16.7*b*.

16.9 See Figure S16.9. From the motion of the hand in Figure 16.9, you see that the end of the string first moved up, then down, and so on. Similarities: Your graph and the snapshot of the wave have identical shapes. Differences: In Figure S16.9, the beginning of the wave motion is on the left at $t = 0$; in the snapshots in Figure 16.9, however, the leading edge of the wave is on the right. So, in addition to your curve having a different scale, it is a mirror image of the wave of Figure 16.9.

Figure S16.9

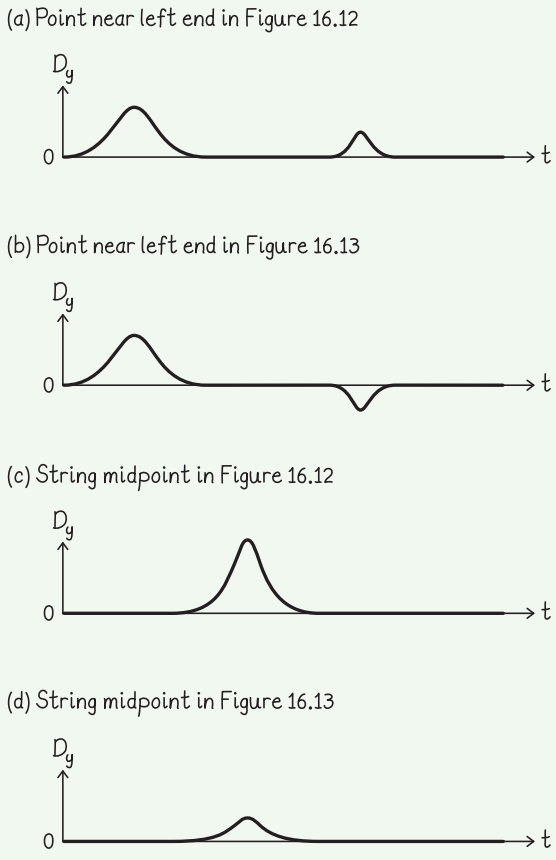


16.10 (*a*) Source: All particles of the string follow the motion of the source. (*b*) Properties of the string. (*c*) Both: The wavelength is determined by the wave speed and the period (see Figure 16.9). The wave speed is determined by the properties of the string, and the period is determined by the source. (*d*) Source: Because each string particle executes the same up-and-down motion as the source, the maximum speed is determined by the source.

16.11 Yes. Imagine putting some object in the path of the pulse. When the pulse displaces the string so that the string hits the object, the object is set in motion, and so the string has transferred momentum to the object.

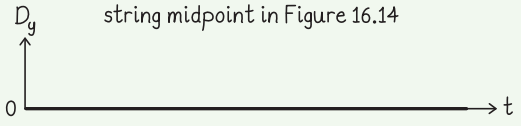
16.12 (*a*) No. To see why, assume for simplicity that the pulses overlap at the midpoint of each string. In Figure 16.12, the maximum displacement at points away from the midpoint where the two pulses overlap is the peak displacement of the larger pulse. In the region of overlap near the midpoint, however, the maximum displacement is greater. In Figure 16.13, the maximum displacement at every point except where the pulses overlap is again the peak displacement of the larger pulse, but the maximum displacement at the midpoint is smaller than this. (*b*) See Figure S16.12*a* and *b* on the next page. In Figure S16.12*a*, the left peak shows how the point near the string's left end is displaced (just after $t = 0$) as the right-moving pulse in Figure 16.12 passes through the point. The right peak shows how, at some later instant, this same point is displaced by the left-moving pulse in Figure 16.12. Figure S16.12*b* shows the same thing for the two pulses in Figure 16.13. (*c*) See Figure S16.12*c* and *d*. Assuming, as in part *a*, that the two pulses overlap at the string midpoint, the Figure S16.12*c* peak is the algebraic sum of the two peaks in Figure 16.12, and the Figure S16.12*d* peak is the algebraic sum of the two peaks in Figure 16.13. (*d*) Comparing Figure S16.12*a* and *c* tells you that a point near the string's left end is displaced twice but the midpoint is displaced only once, justifying the answer no for Figure 16.12. Comparing Figure S16.12*b* and *d* tells you the same thing for Figure 16.13.

Figure S16.12



16.13 (*a*) See Figure S16.13. At the midpoint, the two pulses always cause displacements of equal magnitude in opposite directions. (*b*) The velocity of a point along the string is equal to the derivative of the displacement of that point with respect to time. If the displacement is in the *y* direction, the *y* component of the velocity is $v_y = dD_y/dt$. Because the displacement is the sum of the individual displacements, the velocity is the sum of the individual velocities.

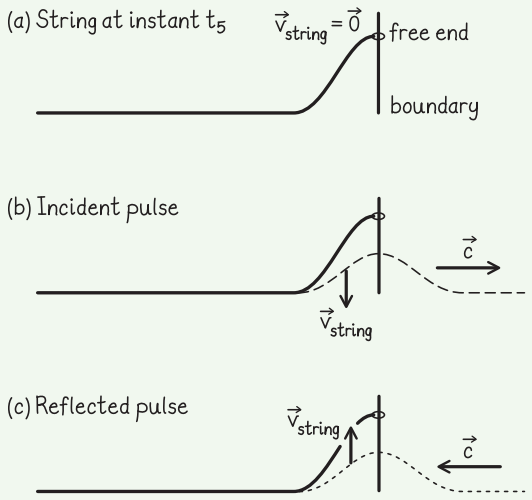
Figure S16.13



16.14 Because the string is straight, its elastic potential energy is zero, which means all the energy is kinetic energy. As Figure 16.16 shows, the pulse is in the process of inverting itself and, although the string is straight, every particle within the wave pulse is in motion.

16.15 (*a*) To see what happens, we construct the pulse shape by the superposition of the incident pulse and an imagined noninverted reflected pulse (Figure S16.15). At t_5 the peaks of both pulses have just reached the free end, and the leading edge of the reflected pulse overlaps the trailing end of the incident pulse. The velocities of particles of the leading edge of the reflected pulse are upward and those on the trailing edge of the incident pulse are downward, so the velocities of particles of the combined pulse are zero (see Checkpoint 16.13). The kinetic energy is zero. (*b*) Let the original pulse have elastic potential energy *U*; then $K = U$ (see Section 16.3) and $E = 2U$. If the string obeys Hooke's law, the potential energy is proportional to the square of the displacement. At t_5, when the displacement at the free end is equal to the sum of the amplitudes of the two pulses (that is, twice as great), the potential energy is four times as great. However, at t_5 the pulse is only half its original width, and so the potential energy is $2U$. (*c*) Yes, $E_1 = E_5 = 2U$.

Figure S16.15



16.16 (*a*) With $\mu_2 = 0$, the boundary becomes a free end, and so the reflected pulse has the same amplitude as the incident pulse and nothing is transmitted to string 2. (*b*) With $\mu_2 = \infty$, the boundary becomes fixed. Again the (inverted) reflected pulse has the same amplitude as the incident pulse and nothing is transmitted to string 2. (*c*) When $\mu_2 = \mu_1$, the strings become one: There is no boundary and so no reflection. The pulse travels unchanged from string 1 to string 2. (*d*) Because the wave speed is greater in string 2, the transmitted pulse is spread out more than the incident pulse (see Section 16.2).

16.17 (*a*) All but (*iii*), which cannot represent a traveling wave because the exponential factor e^{-x} is not a function of $(x \pm ct)$. (*b*) All three, by substituting $(x \pm ct)$ for each x.

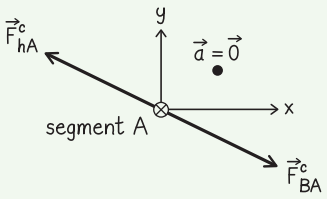
16.18 (*a*) At $t = 0$, the string has zero velocity and so $K = 0$ and all the energy is stored as elastic potential energy. At $t = \frac{1}{8}T$, the string is moving vertically and is not in the equilibrium position, so the energy is distributed between kinetic and elastic potential. At $t = \frac{1}{4}T$, the string is horizontal, so $U = 0$. All the energy is kinetic. (*b*) Yes. Each particle of the string executes simple harmonic motion about its equilibrium position (but different points have different amplitudes—zero at the nodes, maximum at the antinodes). Because the energy of

a simple harmonic oscillator is constant, the energy of any particle or any length of string is also constant. (*c*) No. The amount of energy f_1 carries rightward is equal to the amount f_2 carries leftward, so the combined flow of energy is zero.

16.19 (*a*) No. Suppose one amplitude is twice as large as the other. Half of the large-amplitude wave forms a standing wave with the other wave, but the remaining half remains a traveling wave. The superposition of a standing and a traveling wave is not a standing wave. (*b*) No. Because the wavelengths are not the same, the interference at any position is constructive at some instants and destructive at other instants, unlike the interference illustrated in Figure 16.37, which for a given position is the same at all instants.

16.20 (*a*) See Figure S16.20. Two contact forces act on A: F^c_{hA} and F^c_{BA}. If the gravitational force exerted on A is much smaller than these contact forces, we can ignore it: $\Sigma F_A = 0$. (*b*) The momentum does not change, which makes sense because A moves upward at constant velocity $\vec{v}$. (*c*) Yes, because the scalar product of the force and the force displacement is nonzero and positive: $dW = \vec{F} \cdot d\vec{r}_F$. (*d*) The energy increases because an increasing amount of string moves upward at velocity $\vec{v}$.

Figure S16.20



16.21 Greater (twice as great to be precise) because *two* waves travel away from the point you are shaking, one in each direction along the string. Another way to see this is by looking at the force you must exert to move the string. For a point in the middle, this force is twice as great as the force you must exert on an end because the pieces of string on either side of your hand each exert the same force as one end of the string. So the amount of work you do on the string increases by a factor of 2.

16.22 The frequency f (and therefore the angular frequency ω) is determined by the source that causes the wave. The wave number k is equal to ω divided by the wave speed c, which is determined by the tension and linear mass density of the string.

Chapter 17

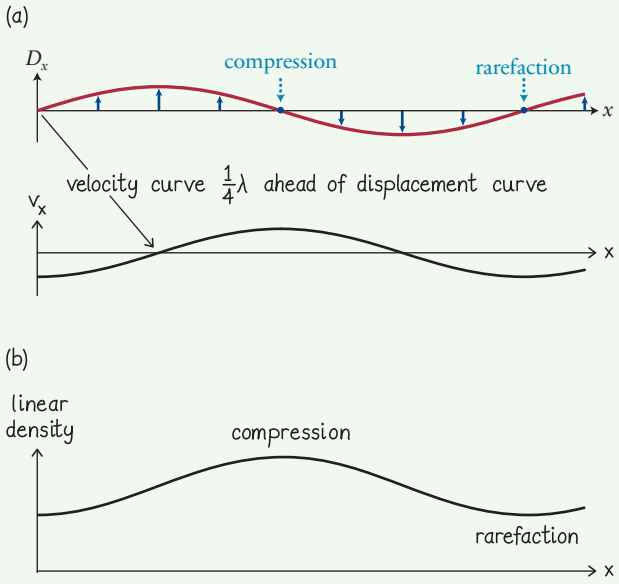
17.1 (*a*) Because the wave speed c is constant, $R_2 = 2R_1$. (*b*) The circumference at R_1 is $2\pi R_1$, making the energy per unit length $E/2\pi R_1$; at R_2 this becomes $E/2\pi R_2 = E/4\pi R_1$. (*c*) The energy varies as $1/r$ because the wavefront circumference increases with r.

17.2 By reducing the amplitude of the source as time passes so that it always matches the decrease in amplitude of the outward traveling wave. Consequently the amplitude of the wave in Figure 17.1, though uniform over *space*, is not constant in time; it uniformly decreases over the entire wave pattern.

17.3 (*a*) Increase. The greater the spring constant, the faster any disturbance is passed from one bead to the next. The effect is the same as when the tension in a string carrying a transverse wave is increased (see Section 16.2). (*b*) Decrease. The greater mass slows down the transmission of the wave just as it does for a transverse pulse on a string of beads.

17.4 (*a*) Just as for a transverse wave, the x component of the velocity of each bead is dD_x/dt. If D_x is a sine function, the x component of the velocity is a cosine function. This means that the velocity wave is always one-quarter wavelength ahead of the displacement wave, as shown in Figure S17.14*a*. (*b*) By comparing Figure 17.9*b* and Figure S17.4*a*, you can see that the linear density is greatest where the velocity is maximum and smallest where the velocity is minimum. Figure S17.4*b* shows a plot of the linear density versus position.

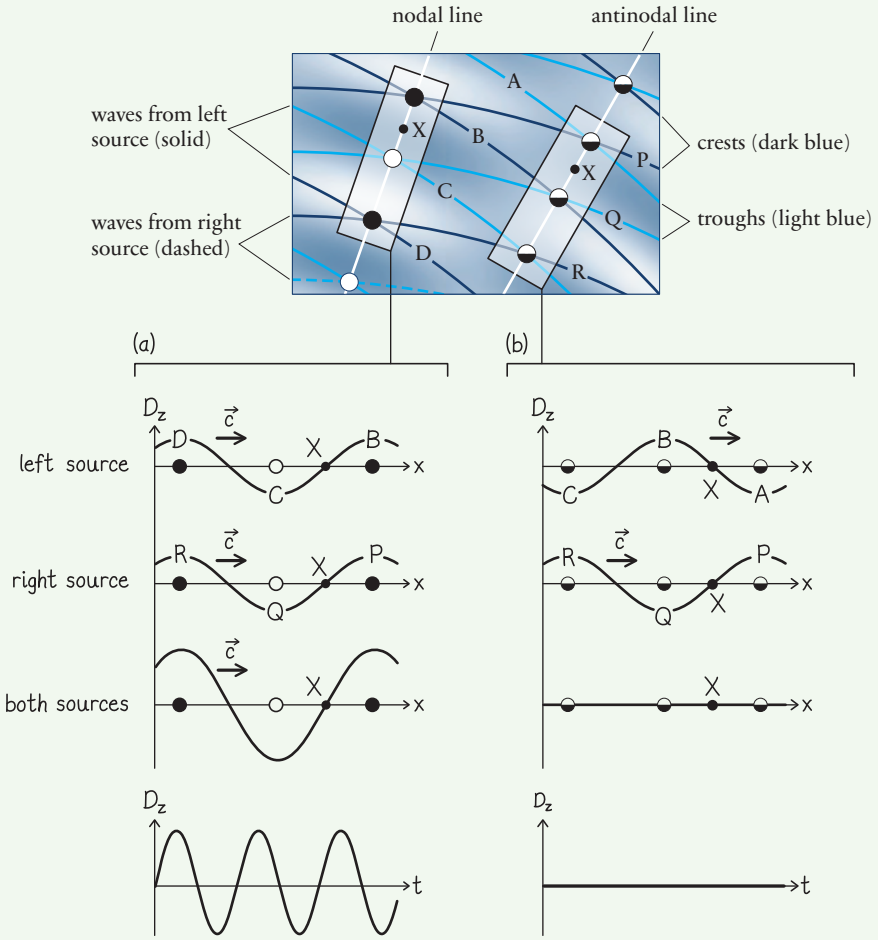
Figure S17.4



17.5 Spherical. The shape of the wavefront is determined by the distance traveled by the wavefront. Because the wave speed is the same in all directions, a given wavefront reaches the same distance from the loudspeaker in all directions. (The sound is louder in the front than in the back because of differences in the wave amplitude, which is greater in front because of the way speakers are designed.)

17.6 (*a*) Zero, because the displacement caused by each wavefront is zero. (*b*) The displacement at that point does not remain zero, however. A quarter period later, two crests (or troughs) meet at that point, and so the displacement at that point varies sinusoidally with an amplitude of 2*A* (Figure S17.6*a* on the next page). (*c*) Along the line that connects the half-filled circles, the waves generated by the two sources are 180° out of phase. Two equal-amplitude sine waves that are 180° out of phase cancel, and so the medium displacement is zero at *all* points along the nodal line, not just the ones marked with half-filled circles. (*d*) See Figure S17.6*b*.

Figure S17.6



17.7 No, because the displacement at this point doesn't remain zero. See Checkpoint 17.6.

17.8 (*a*) $A + A = 2A$ in Figure 17.16*a*; A in Figure 17.16*b*. (*b*) Increase, for two reasons: There is now twice as much energy (two sources), and instead of radiating outward uniformly, the wave energy travels in only certain directions (between nodal lines). (*c*) $E \propto A^2$, and so the energy that passes P increases by a factor of 4. The energy delivered by the sources increases by a factor of 2, while the area over which the energy is spread out is reduced by a factor of 2 (there are just as many antinodal lines as there are nodal lines). These two effects multiply to produce a fourfold increase in the flow of energy at P.

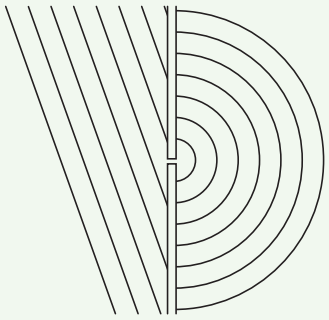
17.9 There are six nodal lines above the horizontal straight line through S_1 and S_2 (and six below also, of course). So $2(d/\lambda) = 6$, $d/\lambda = 3$, $d = 3\lambda$; the distance between the sources is three wavelengths.

17.10 Because along this line, the two waves travel in the same direction and are in phase, which means they reinforce each other everywhere.

17.11 It does not change. Except for a little spreading at the edges, the wavefronts remain straight and parallel. What reaches a 10λ-wide line segment parallel to the row of sources at a distance R also reaches a similar segment at a distance $2R$.

17.12 See Figure S17.12. A narrow gap causes diffraction regardless of the orientation of the incident wavefronts because the incident waves cause the gap to become a point source. (For simplicity the reflections of the planar wavefronts from the back surface of the barrier have been omitted from the drawing.)

Figure S17.12

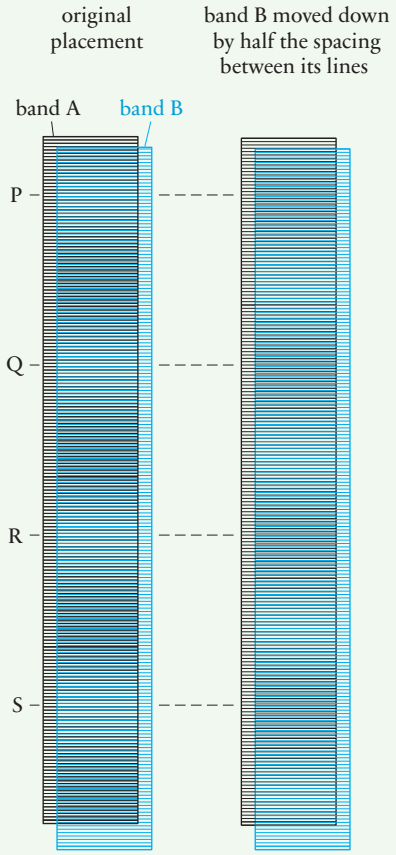


17.13 10 dB corresponds to 1 bel or one order of magnitude (because log 10 = 1). So the intensity must be increased by a factor of 10, which means that ten clarinets must play at the same time.

17.14 (*a*) Three. The A and B lines fall exactly on top of each other at P, but as you move down they get out of step until you reach Q, where they match again. So each band has a whole number of lines between P and Q, but A has one more line. (If they had the same number of lines, each A line would fall on top of a B line. If A had more than one extra line between P and Q, the lines of A would catch up with those of B before Q.) Between Q and S the bands line up twice more, so band A has three additional lines between P and S. (*b*) Points P, Q, R, and S become dark (because the lines of the two

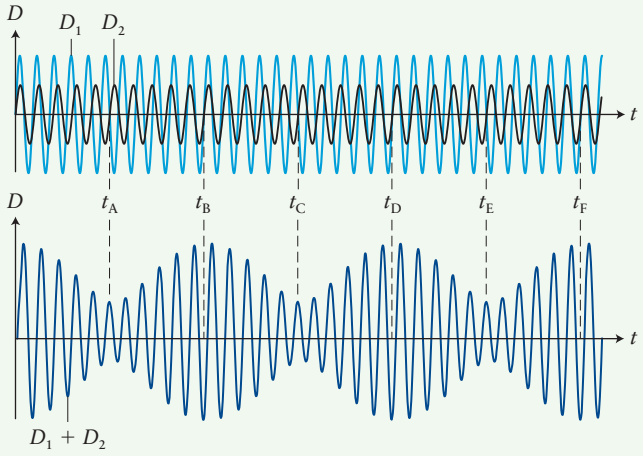
patterns now fall right between each other), while the dark bands turn bright (because the lines now line up; Figure S17.14). Note that the moiré pattern moves down a considerable distance (half the distance between P and Q) even though B moved down only half a line spacing. This effect can be used to measure small displacements.

Figure S17.14



17.15 Yes, but the beats are less pronounced than when the two amplitudes are equal. Let one wave have amplitude A and the other amplitude $2A$. The first wave produces beats with half of the second wave. The other half adds to the beat pattern, resulting in an oscillating wave with a varying amplitude (Figure S17.15), but the amplitude never goes all the way to zero.

Figure S17.15



17.16 Equation 1 in Example 17.8 gives $f_s = f_1(c - v)/c$. Substituting the values for f_1, c, and v yields $f_s = 441$ Hz.

17.17 (*a*) With $v_o = \frac{1}{2}c$ and $v_s = 0$, Eq. 17.20 yields $f_o = \left[\left(c + \frac{1}{2}c\right)/(c - 0)\right]f_s = \frac{3}{2}f_s$. (*b*) With $v_o = 0$ and $v_s = \frac{1}{2}c$, $f_o = \left[(c - 0)/\left(c - \frac{1}{2}c\right)\right]f_s = 2\,f_s$. (*c*) The answers are different because there is a physical difference in the two situations. When the source moves, the wavefronts bunch up or spread out, changing the wavelength. When the observer moves, the wavelength, which is fixed by the wave speed in the stationary medium, does not change. Instead, the moving observer intercepts a different number of wavefronts per unit time. So it is not just the relative motion of source and observer that matters; the fact that their speeds are measured relative to the medium that transmits the wave also matters.

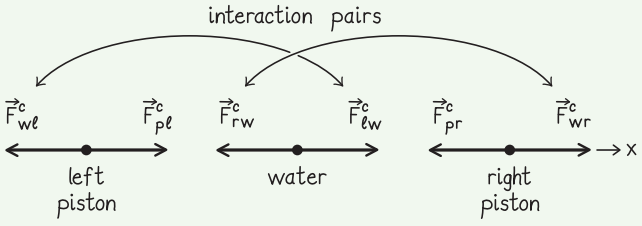
17.18 No. Shock waves are due to the piling up of wavefronts. When the source is stationary and the observer moves, the wavelength is unchanged. Setting $v_o = 2c$ with $v_s = 0$ in Eq. 17.20 yields $f_o = 3f_s$. This reflects the fact that the moving observer intercepts three times as many wavefronts per unit time as an observer at rest.

17.19 The cone angle is about 35°, and sin 35° = 0.57. From Eq. 17.22, you have $c/v_s = 0.57$, $v_s = 1.7c \approx 600$ m/s.

Chapter 18

18.1 (*a*) If the water is to remain at rest, the vector sum of the forces exerted on it must be zero. Because you initially exert a rightward 10-N force on the left piston, you must also exert a leftward 10-N force on the right piston. (*b*) Figure S18.1 shows free-body diagrams for the two pistons and the water. Because of the two interaction pairs marked in the figure, all six forces are equal in magnitude, so $\vec{F}^{c}_{pl} = \vec{F}^{c}_{wr}$. In other words, the force you exert on the left piston is transmitted by the water to the right piston.

Figure S18.1



18.2 (*a*) Yes. The water is at rest, which means the downward force of gravity exerted on it must be balanced by an upward contact force exerted by the bottom of the cup. That the bottom exerts a force on the water tells you that the water must exert a force on the bottom. (*b*) Yes. If the sides of the cup were to disappear, the water would move radially outward. That the water remains in the cup tells you that the sides must exert a force on the water, and consequently the water must exert a force of equal magnitude on the cup in the opposite direction. (*c*) Positive. Because the water pushes outward on the surfaces of the cup, the surfaces must push inward on the water, putting it under compression. By definition, compression means positive pressure.

18.3 If you divide the liquid into three equal layers, each layer is subject to an identical force of gravity. The bottom layer supports the upper two, the middle layer supports only the top one, and the top one supports no other layer. If you make the division even finer, you can see that the pressure decreases linearly with height.

18.4 10 m. The pressure at the surface is 1.0×10^5 N/m^2, and that pressure doubles when the force of gravity exerted on a column of water of height h and cross-sectional area $A = 1$ m^2 has magnitude $F^G = mg = 1.0 \times 10^5$ N. The volume of this column is $V = hA$, its mass is $m = \rho_{\text{water}} V = \rho_{\text{water}}\, hA$, and the magnitude of the force of gravity exerted on it is $mg = \rho_{\text{water}}\, hAg$. Thus $h = mg/(\rho_{\text{water}}\, Ag) = (1.0 \times 10^5 \text{ N})/[(1.0 \times 10^3 \text{ kg/m}^3)(1 \text{ m}^2)(9.8 \text{ m/s}^2)] = 10$ m.

18.5 (*a*) Nothing. Because the pressure in a liquid at rest decreases linearly with height, the pressure difference between the top and bottom of the brick is the same at every water depth. (*b*) Nothing. Pascal's principle tells you that a pressure change applied to a liquid is transmitted undiminished to every part of the liquid. So, increasing the pressure at the water surface increases the pressure at the top and at the bottom of the brick by the same amount, leaving the difference unchanged. (*c*) There is no effect because this pressure at the sides causes equal forces in opposite directions. (*d*) Yes, because the pressure in the atmosphere also decreases with height, although much less than in water. This force is upward.

18.6 The magnitude of the buoyant force exerted on the brick by the water is equal to the magnitude of the force of gravity exerted by Earth on the water displaced by the brick. Because the density of the brick is greater than the density of the water, the mass of the brick must be greater than the mass of the displaced water. Therefore the magnitude $m_{\text{brick}}\, g$ of the force of gravity on the brick is greater than the magnitude $m_{\text{water}}\, g$ of the force of gravity on the displaced water. The magnitude of the upward buoyant force exerted by the water on the brick is thus smaller than the magnitude of the downward force of gravity exerted on it, and so the brick sinks.

18.7 Equal. Because the pan-rock combination floats in both cases, the buoyant force exerted in both cases must be equal to the force of gravity exerted on the displaced water. You also know that the upward buoyant force on the combination must be equal to the downward force of gravity on the combination because the combination remains at rest. The mass of the pan-rock combination doesn't change when the rock is suspended (you can ignore the mass of the string), so the volume of water displaced doesn't change.

18.8 The same. The principle of relativity tells you that the laws of the universe are the same in all inertial reference frames. So, if you view the motion of an object moving at constant velocity through a stationary fluid from the reference frame of the object, you obtain the same situation—a laminar flow of fluid past a stationary object—as in Figure 18.18*a*.

18.9 As you can see in Figure 18.15, streamlines that pass over the roof of the car get closer together, indicating a higher flow speed and therefore lower pressure just above the roof. The air pressure inside the car, however, is unchanged and therefore greater than that outside, making the cloth roof bulge outward.

18.10 According to Pascal's principle, the pressure in the air in the balloon must be the same everywhere.

18.11 (*a*) Zero. The tensile forces due to surface tension are in the plane of the surface and therefore unable to compensate for any pressure difference across the surface when the surface is level. (*b*) Greater. The vector sum of the tensile forces exerted on a segment of the surface points toward the center of the drop. (*c*) It is higher. According to Laplace's law, the smaller radius of curvature of the small drop causes a greater pressure difference across the surface. Because both drops are surrounded by air at atmospheric pressure, the pressure must be greater inside the small drop.

18.12 Because the cylinder has a radius of curvature larger than that of the needle, the upward force exerted by the curved liquid surface on the cylinder is smaller than the upward force exerted on the needle. This smaller upward force is unable to support the (greater) downward force exerted by the cylinder on the surface.

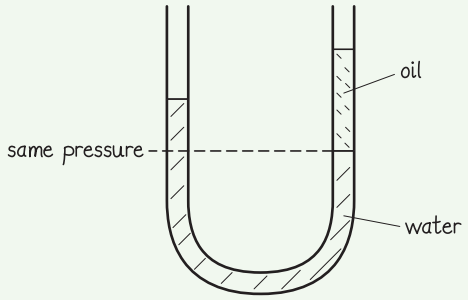
18.13 The magnitude of the force exerted by the air on the cover is $F^c_{\text{ac}} = P_{\text{atm}} A = (101 \text{ kPa})(1000 \text{ Pa}/1 \text{ kPa})(0.28 \text{ m})(0.22 \text{ m}) = 6.2 \times 10^3$ N. The force of gravity exerted on the book is $F^G_{\text{Eb}} = (3.0 \text{ kg})(9.8 \text{ m/s}^2) = 29$ N. The force exerted by the atmosphere on the cover is more than 200 times greater than the gravitational force exerted by Earth on the book.

18.14 As the suction cup hits the ceiling, the bowl of the cup collapses, forcing out the air that was initially in the bowl. Consequently, the pressure in the space between the cup and the ceiling is lower than atmospheric pressure. As long as the seal between the cup edge and the ceiling is maintained, the air in the room underneath the cup exerts an upward force greater than the downward force exerted by the air inside the cup that holds the dart up against the ceiling.

18.15 That the object floats tells you that the buoyant force exerted on it must equal the force of gravity exerted on it, and so equating Eqs. 18.11 and 18.12 gives you $\rho_{\text{o,av}} V_{\text{o}} g = \rho_{\text{water}} V_{\text{disp}} g$. You are given that $V_{\text{disp}} = 0.80 V_{\text{o}}$, and so $\rho_{\text{o,av}} = (\rho_{\text{water}} V_{\text{disp}})/V_{\text{o}} = 0.80 \rho_{\text{water}}$ (see Example 18.2).

18.16 No. The pressure at the oil-water interface in the right leg must be the same as the pressure at a point that is at the same height in the left leg (Figure S18.16). Because the mass density of oil is smaller than that of water, the height of a column of oil required to cause a certain pressure difference is greater than the height of a column of water required to cause the same pressure difference. Therefore the top of the oil column must be higher than the top of the water column.

Figure S18.16



18.17 Because the atmosphere exerts a force on both pistons, it increases the pressure below them. When you add the effect of atmospheric pressure, the pressure in the liquid at the small piston is $P_1 = P_{\text{atm}} + F_1/A_1$ and the pressure at the large piston is $P_2 = P_{\text{atm}} + F_2/A_2$. Ignoring the term $\rho g h$ in Eq. 18.19 again, you have $P_1 \approx P_2$, and so $P_{\text{atm}} + F_1/A_1 = P_{\text{atm}} + F_2/A_2$. The atmospheric pressure cancels out, and you obtain the same result as in Eq. 18.21.

18.18 The continuity equation for an incompressible fluid (which means $\rho_1 = \rho_2$) gives $A_1 v_1 = A_2 v_2$ or $v_1 = v_2 A_2/A_1 = v_2 R_2^2/R_1^2$. Squaring all terms, subtracting v_2^2 from both sides, and dividing through by -1 gives you $v_2^2 - v_1^2 = v_2^2(1 - R_2^4/R_1^4)$. With $R_1 = \frac{1}{2}(0.50 \text{ m})$ and $R_2 = \frac{1}{2}(10.0 \text{ mm})$, you have $(R_2/R_1)^4 = 1.6 \times 10^{-7}$, which is so small that you can ignore this term to get the result $v_2^2 - v_1^2 = v_2^2(1 - R_2^4/R_1^4) \approx v_2^2$. Thus the assumption that $v_1 \approx 0$ in Example 18.9 is excellent, and you can use the result in that example: $v_2 = \sqrt{2gd} = 2.4$ m/s.

18.19 A soap bubble consists of a thin film of soap solution surrounding a volume of air. Because the film has two surfaces, the force due to surface tension gains an additional factor 2. Thus Eq. 18.51 becomes $P_{\text{in}} - P_{\text{out}} = 4\gamma/R$.

Chapter 19

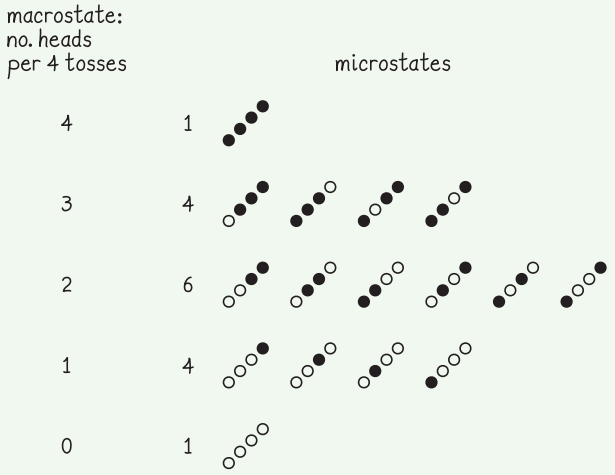
19.1 (*a*) Collisions with the leading edge of the pendulum slow it down, and collisions with the trailing edge speed it up. Because of the motion of the pendulum, nitrogen molecules that collide with the leading edge move faster relative to the pendulum than nitrogen molecules that collide with the trailing edge. Therefore there are more collisions with the leading edge than with the trailing edge, and the pendulum slows down. (*b*) When the pendulum is at rest, the average speed of the nitrogen molecules that hit the left edge is the same as the average speed of the molecules that hit the right edge, which means that the momentum transfers due to the collisions tend to balance out.

19.2 Each segment of the zigzag path would decrease in length, so the Brownian motion would be less pronounced and the maximum displacement of the grain from the initial position would be smaller.

19.3 7.

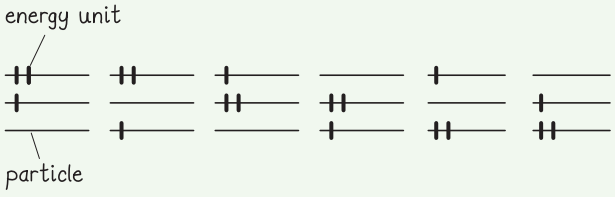
19.4 See Figure S19.4. (*a*) Five macrostates. (*b*) 16 basic states. (*c*) $4/16 = 1/4$. (*d*) The most likely macrostate is two heads because there are six basic states associated with this macrostate.

Figure S19.4



19.5 See Figure S19.5.

Figure S19.5



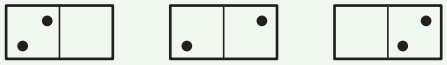
19.6 (*a*) $1/84$, which is around 1%. (*b*) $\frac{28}{84} + \frac{21}{84} = \frac{49}{84}$, which is more than 50%.

19.7 (*a*) The pendulum has more than 50% of the energy when it has four, five, or six units of energy . There are $6 + 3 + 1 = 10$ basic states associated with these macrostates. The probability that the pendulum has more than 50% of the energy is then $10/84 = 0.12$. (*b*) Figure 19.7 shows that as the number of particles increases, the relative probability of any given fraction of energy being in the pendulum decreases.

19.8 (*a*) The average kinetic energies are equal because the energy of the gas is equally distributed among all the particles. (*b*) Because the two particles have the same average kinetic energy, $\frac{1}{2} m_1 v_1^2 = \frac{1}{2} m_2 v_2^2$. Therefore $v_1 = (m_2/m_1)^{1/2} v_2 = 2^{1/2} v_2$. Particle 1 has a greater average speed by a factor of $2^{1/2}$.

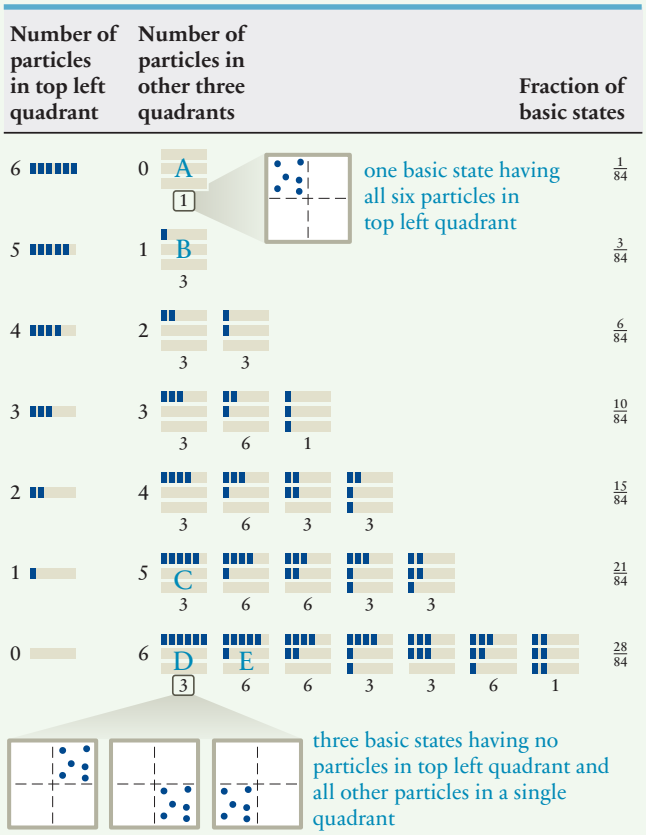
19.9 There are three ways of placing the two indistinguishable particles in the two compartments (Figure S19.9). The probability of finding one particle in each side is $1/3$.

Figure S19.9



19.10 (*a*) Four basic states have all six particles in one compartment. They originate from different macrostates. Distribution A in Figure S19.10 has six particles in the top left compartment and is a macrostate. Distribution D is part of a different macrostate: no particles in the top left compartment. The three basic states in distribution D correspond to six particles in any of the other three compartments. There are 84 basic states, all equally probable, and so the probability of finding six particles in any compartment is $4/84 = 0.048$. (*b*) There are 12 basic states in which one compartment contains five particles: three for distribution B in Figure S19.10, three for distribution C, and six for distribution E. The probability of finding five particles in any compartment is thus $12/84 = 0.143$. Note that distribution E has six associated basic states, corresponding to the six ways one particle can be placed in one of three containers and five particles can be placed in another of the three containers.

Figure S19.10



19.11 The partition moves until the particles are uniformly distributed over the space enclosed by the two compartments. Because there are twice the number of particles in the left compartment, the partition moves to the right until the volume of the left compartment is twice the volume of the right compartment.

19.12 (*a*) All six energy units are initially in the pendulum, which means each collision moves one energy unit from the pendulum to one of the particles. The system then is in one of the three basic states depicted in the second of row of Figure 19.4. (*b*) Now there are three possibilities for how energy is transferred in any given collision: from the particle that initially has the one unit to either of the two other particles, from the particle that initially has the one unit to the pendulum, or from the pendulum to any one of the particles. When a unit is transferred from one particle to another, the amount of energy in the pendulum does not change, and so the macrostate does not change (that macrostate is represented in the second row in Figure 19.4). In a collision between the pendulum and a particle, the pendulum can either gain or lose a unit of energy, and so the system ends up in the macrostates represented by the top and third rows in Figure 19.4. So, after one collision the system can be in any one of the top three macrostates in Figure 19.4. (*c*) The number of basic states in the top three macrostates is $1 + 3 + 6 = 10$.

19.13 (*a*) There are three accessible macrostates: three units in pendulum, two units in pendulum, one unit in pendulum. (*b*) To find the probability for each macrostate, you need to consider the number of basic states representing each: 10, 15, and 21, respectively, for a total of 46. The probabilities are thus $10/46$, $15/46$, and $21/46$.

19.14 (*a*) There are 21 ways that two energy units can be distributed among the six particles. Label the particles 1 through 6. Place one energy unit in particle 1. The second energy unit can be placed in any of the six particles, resulting in six basic states. Now place one energy unit in particle 2. The second energy unit can be placed in the five particles 2 through 6 but not in particle 1 because we already counted that basic state. Repeat this process for particles 3 through 6. The number of basic states is then $6 + 5 + 4 + 3 + 2 + 1 = 21$.

19.15 (*a*) There are 14 particles in A, so the average energy per particle in A is $1/14 = 0.07$ energy unit. The average energy per particle in B is $9/6 = 1.5$ energy units. (*b*) At equilibrium there are seven energy units in A, corresponding to an average energy per particle of $7/14 = 0.5$ energy unit. The average energy per particle in B is $3/6 = 0.5$ energy unit.

19.16 (*a*) Over time the ice cube heats up and melts, and the water cools. Because the system irreversibly evolves toward equilibrium, Ω must be increasing. (*b*) Because everything—cup, tea, air—is in equilibrium, Ω is at its maximum value and not changing. (*c*) Over time the oil separates and floats to the top. Because this separation constitutes an irreversible process (the oil doesn't spontaneously mix back into the water), Ω must be increasing. (*d*) Germination is an irreversible process, so Ω is increasing. (*e*) As far as we know, the universe evolves irreversibly, so Ω is increasing.

19.17 (*a*) The compartment size decreasing by a factor of 10 means that $M = 1000$. The number of basic states is then $\Omega_f = M^N = 1000^{10} = 10^{30}$. The initial value is $\Omega_i = 10^{20}$ (see Example 19.8), so the change in the number of basic states is $10^{30} - 10^{20}$, which for all practical purposes is equal to 10^{30}. This represents an enormous increase in the number of basic states. (*b*) $\ln \Omega_f = \ln (10^{30}) = 69$. The initial value is $\ln \Omega_i = 46$ (see Example 19.8), so the change is $69 - 46 = 23$.

19.18 (*a*) According to Eq. 19.4, an increase in entropy means the number of basic states has increased. (*b*) For an expanding gas, $V_f > V_i$, and so the $N \ln (V_f/V_i)$ term in Eq. 19.8 is positive, as expected for a system evolving toward equilibrium. (*c*) If the gas were to contract into a subvolume, $V_f < V_i$, and so the $N \ln (V_f/V_i)$ term is negative, meaning a decrease in entropy. Such a decrease would violate the entropy law.

19.19 Even though entropy depends on volume, the entropy change depends on a ratio of volumes (Eq. 19.8). Because the volume of a compartment is a fraction of the volume V of the box, V cancels when we take the ratio of volumes for each compartment.

19.20 (*a*) Adding up the x components of the velocities and dividing by 5 gives an x component of the average velocity of 0. (*b*) Adding up the speeds (the absolute values of the x components of the velocities) and dividing by 5 gives an average speed of 2.6 m/s. (*c*) Adding up the squares of the speeds, dividing by 5, and then taking the square root gives an rms speed of 3.0 m/s. Note that the rms speed is not equal to the average speed of the atoms: in a gas, the former is typically about 10% greater than the latter.

19.21 Putting all the numerical factors in the third term of Eq. 19.29 together gives

$$\left[2a\left(\frac{2E_{th}}{mN}\right)^{1/2}\right]^3 = \left[\left(\frac{8a^2E_{th}}{mN}\right)^{1/2}\right]^3 = \left(\frac{8a^2}{mN}\right)^{3/2}(E_{th})^{3/2}.$$

19.22 (*a*) At thermal equilibrium there is equipartition of energy, and so, on average, the kinetic energy of all the atoms is the same. To the left of the dashed line, S_A is smaller and S_B is greater than at thermal equilibrium. According to Eq. 19.32, $E_{th,A}$ is therefore smaller than the equilibrium value and $E_{th,B}$ is greater than the equilibrium value. Consequently, to the left of the dashed line the average kinetic energy of the atoms in A is smaller than that in B. (*b*) At thermal equilibrium, both gases are equally hot. To the left of the dashed line gas B is hotter because the average kinetic energy of the atoms is greater. (*c*) At thermal equilibrium, $dS_A/dE_{th,A} = dS_B/dE_{th,B} = -dS_B/dE_{th,A}$ (Eq. 19.37). Figure 19.24 shows that the magnitude of the slope of curve S_A increases and the magnitude of the slope of curve S_B decreases when going left from the equilibrium point. Therefore, to the left of the equilibrium point $dS_A/dE_{th,A} > dS_B/dE_{th,B}$.

19.23 From Eq. 19.42 we have $E_{th,B}/E_{th,A} = N_B/N_A = 6/14 = 3/7$, and so 30% of the energy is in compartment B and 70% is in compartment A.

19.24 (*a*) Only the component of the momentum perpendicular to the wall changes. The atom's final momentum is away from the wall, and so the change in momentum is perpendicular to the wall and directed toward the interior of the container. (*b*) Because the system is isolated, the direction of the wall's change in momentum is opposite the direction of the atom's change in momentum. (*c*) As you saw in Section 8.7, the force exerted on an object is proportional to the object's change in momentum. The force exerted by the wall on the atom is therefore perpendicular to the wall and directed toward the interior of the container. The force exerted by the atom on the wall points in the opposite direction.

19.25 In deriving Eq. 19.20 we assumed that the system is in thermal equilibrium (see the beginning of Section 19.5), and so $T_A = T_B$. Multiplying the left and right sides of Eq. 19.20 by $k_B T_A$ gives $N_A k_B T_A/V_A = N_B k_B T_B/V_B$, or using the ideal gas law, $P_A = P_B$. In equilibrium, the pressure is the same on both sides of the partition, as we would expect.

19.26 (*a*) Neither. Because the pressures and volumes are equal, the thermal energies of the two gases are equal. (*b*) Neither. Because the two containers have the same number of atoms, the average kinetic energy is the same for all atoms. (*c*) Argon. The rms speed is proportional to the square root of the average kinetic energy divided by the mass (Eq. 19.53). Although the average kinetic energies of the two gases are equal, a helium atom has a smaller mass than an argon atom, so the rms speed of a helium atom is greater than that of an argon atom.

19.27 (*a*) No, because the system is not closed; as the gas cools, thermal energy leaves the system. The entropy law applies only to closed systems. (*b*) From Example 19.12*a*, we have $T_f = 145$ K, so $\Delta T = 145\text{ K} - 290\text{ K} = -145$ K, and so, from Eq. 19.50: $\Delta E_{th} = \frac{3}{2}Nk_B\Delta T = \frac{3}{2}(10^{23})(1.38 \times 10^{-23}\text{ J/K})(-145\text{ K}) = -300$ J.

19.28 Decrease. From Eq. 19.61, because $T_f = T_i/2$ and $V_f = 2V_i$, we have

$$\Delta S = \tfrac{3}{2} N \ln \tfrac{1}{2} + N \ln 2.$$

Because $\ln \frac{1}{2} = -\ln 2$, this can be written as

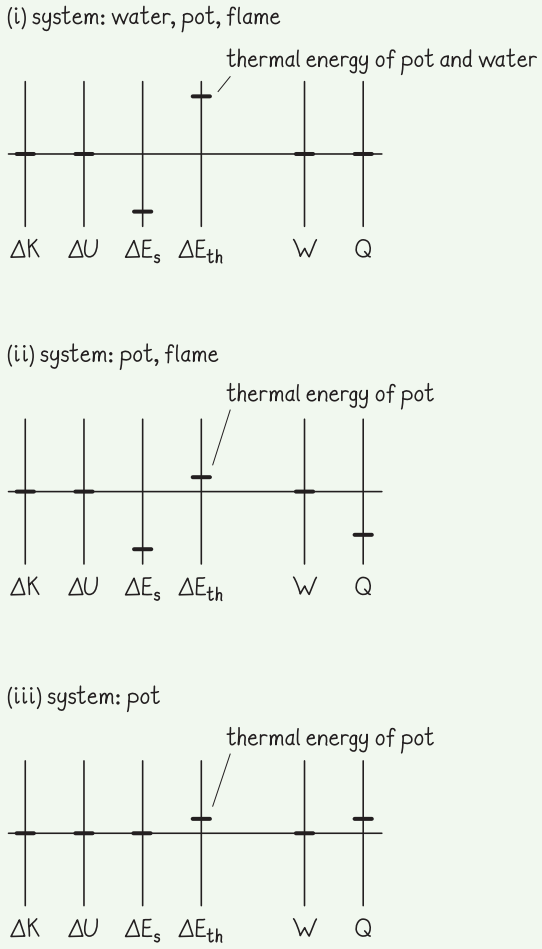
$$\Delta S = -\tfrac{3}{2} N \ln 2 + N \ln 2 = -\tfrac{1}{2} N \ln 2 < 0,$$

and so the entropy decreases. When the temperature of the system is halved, the energy of the system decreases, and so the system is not closed.

Chapter 20

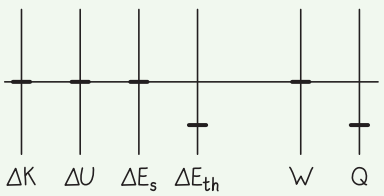
20.1 (*a*) Joules. (*b*) See Figure S20.1.

Figure S20.1



20.2 See Figure S20.2.

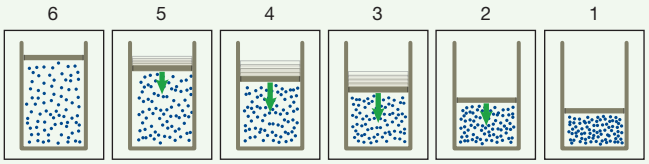
Figure S20.2



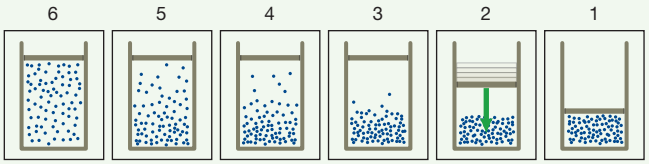
20.3 See Figure S20.3. The quasistatic process is possible because it depicts a succession of equilibrium states, all of which show an equipartition of space. The non-quasistatic process is not possible: As the first four frames show, it would require the gas molecules to "pull away" spontaneously from the piston, before the piston begins moving down in frame 2.

Figure S20.3

(*a*) Reverse quasistatic process



(*b*) Reverse non-quasistatic process



20.4 During a quasistatic process the system must remain near equilibrium at all instants during the process. (*a*) Not quasistatic; just after the balloon is popped, the pressure inside the container is not the same everywhere and so the system is not in equilibrium. (*b*) Quasistatic; even though the temperature of the coffee changes, the change is slow. The coffee therefore has a well-defined uniform temperature and is in equilibrium throughout the cooling process. (*c*) Quasistatic; as the gas pushes the piston outward, the pressure in the gas remains uniform and so the gas is in equilibrium. (*d*) Not quasistatic; the temperature of the ice cube and the hot water are different, and so the system is not in equilibrium.

20.5 (*a*) Using Eq. 2 from the "Temperature measurement scales" box, you find for $T_F = 60\ ^\circ\text{F}$:

$$T_C = \frac{5\ ^\circ\text{C}}{9\ ^\circ\text{F}} (60\ ^\circ\text{F} - 32\ ^\circ\text{F}) = 16\ ^\circ\text{C}.$$

Likewise, for $T_F = 80\ ^\circ\text{F}$, you obtain $T_C = 27\ ^\circ\text{C}$, making the range 16 °C to 27 °C. By adding 273.15 to these values, you get the corresponding range in kelvins: 289 K to 300 K.

(b) Equation 1 from the box gives you

$$T_C = \frac{1\ ^\circ\text{C}}{1\ \text{K}} (2800\ \text{K}) - 273.15\ ^\circ\text{C} = 2500\ ^\circ\text{C}.$$

The expression I used in the solution to Exercise 20.2*b* gives

$$T_F = \frac{9\ ^\circ\text{C}}{5\ ^\circ\text{F}} \left[\frac{1\ ^\circ\text{C}}{1\ \text{K}} (2800\ \text{K}) - 273.15\ ^\circ\text{C} \right] + 32\ ^\circ\text{F} = 4600\ ^\circ\text{F}.$$

20.6 (*a*) There are many more water molecules in the pool than in the glass. Thermal energy is proportional to number of molecules, and so the pool has greater thermal energy. (*b*) The heat capacity of

the pool is greater because the amount of thermal energy needed to increase the temperature of all the water in the pool by 1 K is greater than the amount needed to increase the temperature of all the water in the glass by 1 K. (*c*) The glass of warm water because thermal energy is proportional to temperature.

20.7 (*a*) All three. Unlike a hydrogen molecule, an ammonia molecule is not linear, so its rotational inertia is nonzero along the three orthogonal axes. (*b*) Because there are six degrees of freedom, three translational and three rotational, the thermal energy per molecule is $6(\frac{1}{2}k_BT) = 3k_BT$. (*c*) If you raise the temperature by $\Delta T = 1$ K, the increase in thermal energy per molecule is $\Delta E_{th} = 3k_B\Delta T$. The heat capacity per molecule is thus $\Delta E_{th}/\Delta T = 3k_B$. (*d*) It's within 10%. The experimental value listed in Table 20.3 is $3.37k_B$.

20.8 Less than. More energy must be transferred to the diatomic gas because it has more degrees of freedom, and so its heat capacity per particle is greater.

20.9 $P_1 > P_2$, $V_1 < V_2$, and $T_1 = T_2$ because the two states are on an isotherm. The entropy dependence on temperature and volume is given by Eq. 19.60, $S \propto \ln T^{3/2}V$. The temperatures are the same in the two states, and so the differences in entropy depend only on volumes. That $V_1 < V_2$ tells you that $S_1 < S_2$. Because $E_{th} \propto T$, the thermal energies of the two states are the same.

20.10 (*a*) The change in entropy is the same for the two processes (the same initial and final states and therefore the same initial and final temperatures and volumes). (*b*) No. Because the two processes have the same initial and final states, the change in temperature is the same for both. Hence the change in thermal energy ΔE_{th} is the same. This change is the sum of the work done on the gas and the energy transferred thermally to it: $\Delta E_{th} = W + Q$. Because the work done on the gas (the area under the path) is different for the two processes, the amount of energy transferred thermally must also be different.

20.11 See Table S20.11. The work done on the gas is positive during compression, negative during expansion, and zero for any isochoric process.

The thermal energy is proportional to the temperature (see Eq. 19.50). As you can see from Figure 20.25, the isentropic expansion brings the system to a lower isotherm, and so the thermal energy decreases; in the isochoric process, the temperature increases, and so $\Delta E_{th} > 0$; the isobaric process decreases the temperature, and so $\Delta E_{th} < 0$; and the isothermal compression doesn't change the temperature, so $\Delta E_{th} = 0$.

The isentropic process is adiabatic, so $Q = 0$. For the other processes, the sign of Q can be determined from ΔE_{th} and W, because the sum of the work done on the gas and the energy transferred thermally to it equals the change in thermal energy.

For the isentropic process the change in entropy is zero by definition. As you can see from Figure 20.25, the isochoric process brings the system to a higher isentrope and so $\Delta S > 0$. The isobaric and isothermal compressions bring the system to a lower isentrope, so $\Delta S < 0$. (Because the system is not closed, the entropy can decrease.)

20.12 Table 20.3 and Figure 20.15 tell you that at $T = 300$ K, the heat capacity per particle for hydrogen is $2.5k_B$ and so it has five degrees of freedom (three translational and two rotational). So, if we assume the gas can be treated as ideal, the thermal energy is $E_{th} = \frac{5}{2}Nk_BT$ (approximately N times 10^{-20} J).

20.13 Process A. The vertical legs for both A and B are at constant volume, and Eq. 20.9 tells you that no work is done on the gas. In process A, the gas undergoes an isobaric decrease in volume at a higher pressure than process B. The volume change is the same in both cases. So, by Eq. 20.9 (which has ΔV no longer equal to zero), the work done on the gas during the isobaric process in process A is greater than in process B. (Note that the area under path A in Figure 20.28 is larger than the area under path B and the work done on the gas is positive, thus confirming our answer.)

20.14 (*a*) Process C. Because the initial and final states lie on an isotherm, $\Delta E_{th} = 0$ for all three processes. The work done on the system in each process is negative and equal in magnitude to the area under the path. Because $\Delta E_{th} = 0$, $Q = -W$ and so Q is equal to the area under the path, which is largest for process C and (*b*) smallest for process A.

20.15 Yes. If $\Delta T = 0$, then the temperature doesn't change, and if the temperature doesn't change, then Eq. 20.5 tells us that E_{th} doesn't change either, so $\Delta E_{th} = 0$.

20.16 Greater. Because $PV = Nk_BT$ and Nk_B is constant, the product PV is proportional to the temperature T of the state. The final temperature T_f is greater than the initial temperature T_i, and so $P_fV_f > P_iV_i$.

20.17 The isobaric process. Both processes have the same ΔE_{th} because the temperature change is the same. No work is done on the gas during the isochoric process, which means that in this process the energy transferred thermally goes entirely into raising the temperature. During the isobaric process, more energy must be transferred thermally into the gas during the expansion because the work done on the gas is negative.

20.18 Because the initial and final states lie on the same isotherm, the thermal energy of the gas doesn't change. From this you know that $\Delta E_{th} = W + Q = 0$ and so $W = -Q$. The work done on the gas is positive for both processes ($V_f < V_i$), and so the energy transferred thermally must be negative. (*a*) The work done on the gas is equal to the area under the path. From the paths in Figure 20.37 you see that the rectangular area under path A is larger than the area under path B, which means that the work done on the gas in process A is greater than the work done on the gas in process B. (*b*) Because

Table S20.11 Checkpoint 20.11

Process	ΔE_{th}	W	Q	ΔS
isentropic expansion	negative	negative	0	0
isochoric pressure increase	positive	0	positive	positive
isobaric volume decrease	negative	positive	negative	negative
isothermal compression	0	positive	negative	negative

W is greater for process A, the magnitude of Q must be greater for process A.

20.19 No. Equation 20.36 gives the change in entropy between two equilibrium states that are at the same temperature. During an isochoric process, the temperature changes, so we cannot use Eq. 20.36.

20.20 The same. Equation 20.34 is independent of the process. The entropy change depends on only the temperatures and volumes of the initial and final states. If the two combinations of processes have the same initial and final states, the change in entropy is the same in the two cases.

20.21 See Table S20.21. ΔS is positive when a process ends on a higher isentrope. To determine the sign of Q, we approximate the ideal gas process by a series of infinitesimal isentropic and isothermal segments. (We can do this for any process.) Along the isentropic segments, $Q = 0$. Along the isothermal segments, Q has the same sign as ΔS, which is positive. Therefore for any process for which $\Delta S > 0$, we know that Q must be positive too. The change in the thermal energy is the sum $Q + W$. Because isentropes are steeper than isotherms, ending on a higher isentrope while compressing the gas means ending on a higher isotherm, and so the temperature T and the thermal energy both have to increase ($\Delta E_{th} > 0$). When the gas expands, its entropy can increase while its temperature and thermal energy increase, decrease, or stay the same. (*a*) $W < 0$ for expansion. (*b*) $W > 0$ for compression.

Table S20.21 Checkpoint 20.21

	Expansion	Compression
W	negative	positive
Q	positive	positive
ΔE_{th}	positive, negative, or zero	positive
ΔS	positive	positive

20.22 We have to calculate entropy changes associated with two processes: the cooling down of the water from 20 °C to 0.0 °C and the phase transition from water to ice. The entropy change during the cooling is given by Eq. 20.53. Substituting the specific heat capacity for water from Table 20.2, we obtain

$$\Delta S = \frac{(0.010\text{ kg})(4181\text{ J/kg}\cdot\text{K})}{1.38\times10^{-23}\text{ J/K}}\ln\left(\frac{273}{293}\right) = -2.1\times10^{23}.$$

For the phase transition, we use Eq. 20.58 and substitute the specific transformation energy for water from Table 20.4:

$$\Delta S = -\frac{(0.010\text{ kg})(3.34\times10^{5}\text{ J/kg})}{(1.38\times10^{-23}\text{ J/K})(273\text{ K})} = -8.9\times10^{23}.$$

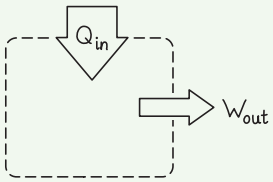
The total entropy change is thus -11×10^{23}.

Chapter 21

21.1 (*a*) There is a change in the state of motion of the puck, which decreases its kinetic energy. Because the temperature of the puck and surface increases, their thermal energy increases. This is an irreversible process, so the entropy increases. (*b*) The spring relaxes from the compressed state, and the cart changes its state of motion. The configurational potential energy of the spring-cart system decreases, and the kinetic energy of the cart increases. The process is reversible, so the entropy does not change. (*c*) A chemical reaction changes the state of the explosives, releasing source energy. The projectile changes its state of motion, increasing its kinetic energy. The temperature and thermal energy of the cannon and projectile increase. Because the process is irreversible, the entropy increases. (*d*) The temperature of the hot object decreases, and therefore its thermal energy decreases. The temperature of the cold object increases, and therefore its thermal energy increases. This process is irreversible, so the entropy increases.

21.2 (*a*) See Figure S21.2. Energy is transferred thermally to the gas. This results in an increase in both the temperature and the volume of the gas. The expanding gas does work on the piston. (*b*) Both change. The temperature of the gas increases, so its thermal energy increases. The entropy increases because both the temperature and volume of the gas increase (Eq. 19.61).

Figure S21.2



21.3 (*a*) Pushing the piston back to its original position requires us to do as much work on the device as the device has delivered to us (Eq. 20.30). Therefore once the piston is back to its original position $W = 0$, and because $\Delta E = 0$, $Q = 0$ too. No energy has been converted. (*b*) The volume and temperature of the gas are the same before expansion and after compression, so by Eq. 19.61 the entropy stays the same. (*c*) Yes, because the device keeps returning to its initial state.

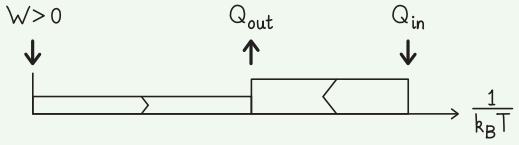
21.4 (*a*) $\Delta E_R > 0$ (energy increases), $\Delta S_R > 0$ (entropy increases with energy), $\Delta S_R/\Delta E_R > 0$. (*b*) $\Delta E_R < 0$, $\Delta S_R < 0$, $\Delta S_R/\Delta E_R > 0$.

21.5 (*a*) Energy quality decreases because $\Delta S > 0$ means that energy is transferred to the right on the dS/dE axis, as in Figure 21.16*b*. (*b*) Energy quality remains the same because it remains at the same value of dS/dE.

21.6 (*b*), (*c*), and (*f*).

21.7 See Figure S21.7.

Figure S21.7



21.8 Because the process depicted in Figure 21.22*a* is reversible, $\Delta S_{env} = 0$, and so the two rectangles must have equal areas. The rectangle on the right is less long and therefore must have a greater height.

21.9 (*a*) Because you increase T_{in}, the label Q_{in} and the point where the two rectangles meet both move to the left. The length of the upgrade rectangle decreases, which means the magnitude of ΔS corresponding to the upgrade decreases. (*b*) Because the process is reversible, the magnitude of ΔS associated with the degradation also decreases. (*c*) Because the length of the degrade rectangle increases, the height must decrease; therefore Q_{out} decreases. That means, because $W = Q_{out} - Q_{in}$ and W is constant, Q_{in} decreases.

21.10 Lowering T_{out} requires a smaller thermal input of energy for the same work done on the environment, so the efficiency increases.

21.11 (*a*) In Figure 21.26, the Q_{in} arrow is roughly twice as wide as the W_{in} arrow. Because $W = W_{in}$, the coefficient of performance of cooling is $Q_{in}/W = 2$. (*b*) Because $Q_{out} = Q_{in} + W_{in} = 3W_{in} = 3W$, the coefficient of performance of heating is $Q_{out}/W = 3W/W = 3$.

21.12 (*a*) Now the area under the expansion path $1 \rightarrow 2$ is greater in magnitude than the area under the compression path $2 \rightarrow 1$, which means the negative work done on the system is greater than the positive work done on the system. So the work done on the system during the cycle is negative, and the magnitude is equal to the shaded area in Figure 21.28*d*. (*b*) The sign of the work done on the environment is the opposite of the sign of the work done on the system and so is positive. The magnitude is the same as in part *a*.

21.13 Although the transfer from low temperature to high temperature is an upgrade of the energy and therefore a decrease in entropy, the entropy increase from the degradation of the mechanical energy input is greater than the decrease.

21.14 Zero. The law of energy conservation requires that the energy of the closed system comprising the device and its environment cannot change. Because the energy of the steady device is not changing, $\Delta E = 0$ (Eq. 21.1), the energy of the environment cannot change either.

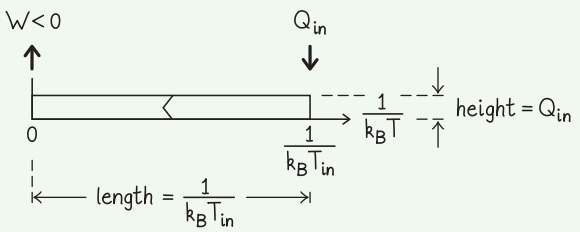
21.15 (*a*) Both negative. Because energy is transferred thermally out of the bar, the energy of the bar decreases and so its entropy decreases. (*b*) Energy is transferred thermally to the environment, and S_{env} increases. (*c*) The equilibration is irreversible, so ΔS_{env} must be greater than ΔS_{bar}, which means $S_{bar} + S_{env}$ increases.

21.16 (*a*) See Figure S21.16. (*b*) Because the transfer is to the left, the entropy change is negative. The height of the rectangle is Q_{in}. (Because $W = -Q_{in}$ you could also write the height of the rectangle as $-W > 0$.) The rectangle stretches from the origin (0 on the $1/k_BT$ scale) to $1/k_BT_{in}$, so its length is $1/k_BT_{in}$. The entropy change is thus

$$\Delta S_{env} = -(Q_{in})\left(\frac{1}{k_BT_{in}}\right) = \frac{-Q_{in}}{k_BT_{in}}.$$

This process is not allowed by the entropy law because $\Delta S_{env} < 0$.

Figure S21.16



21.17 Yes, in theory. According to Eq. 21.20, Q_{out} goes to zero if T_{out} goes to zero. As T_{out} goes to zero, the length of the degrade rectangle becomes infinite, but the area of the rectangle can remain constant. Unfortunately there are no thermal reservoirs at $T = 0$.

21.18 (*a*) $\eta_{max} = (299\text{ K} - 221\text{ K})/299\text{ K} = 0.261$. (*b*) $\eta_{max} = (300\text{ K} - 221\text{ K})/300 = 0.263$, so the efficiency increases by 0.002.

21.19 The blanket converts mechanical energy entirely to thermal energy, and so from Eq. 21.24, $COP_{heating} = Q_{out}/W = 1$, which is the smallest possible value.

21.20 (*a*) Equation 21.28 gives $COP_{heating} = COP_{cooling} + 1$. You know from Example 21.7 that $COP_{cooling}$ is 17. Therefore $COP_{heating}$ is 18. (*b*) You can use Eq. 21.24 to get the delivery rate: $dW/dt = (dQ_{out}/dt)/COP_{heating} = (500\text{ W})/18 = 28\text{ W}$.

21.21 As the gas goes from state 1 to state 2 the entropy change of the gas is (Table 21.1):

$$\Delta S_{1\rightarrow 2} = N \ln\left(\frac{V_2}{V_1}\right).$$

Substituting the values given then yields

$$\Delta S_{1\rightarrow 2} = 5.10 \times 10^{23}.$$

During the two isentropic legs, the entropy doesn't change:

$$\Delta S_{2\rightarrow 3} = 0$$

$$\Delta S_{4\rightarrow 1} = 0.$$

The entropy change over the entire cycle is zero, so the entropy change as the gas goes from state 3 to state 4 must be the negative of the entropy change as the gas goes from state 1 to state 2:

$$\Delta S_{3\rightarrow 4} = -\Delta S_{1\rightarrow 2} = -5.10 \times 10^{23}.$$

21.22 (*a*) The efficiency of the jet engine is given by Eq. 21.44. Substituting T_1 and T_4 from Example 21.9, we get $\eta = 1 - (T_4/T_1) = 1 - (288\text{ K})/(760\text{ K}) = 0.62$. (*b*) Equation 21.36 gives the Carnot efficiency: $\eta = (1520\text{ K} - 288\text{ K})/(1520\text{ K}) = 0.81$, which is significantly greater than the 0.62 efficiency of the engine running on the Brayton cycle. (Remember, however, that the Carnot cycle operates very slowly, which means the power is very low.)

Chapter 22

22.1 (*a*) You should observe that the tape and battery attract each other. (If the battery repels the tape, you can get rid of the repulsion by wiping the entire battery surface with your hand.) It makes no difference how you orient the battery or whether the battery is fresh or spent. Any wooden object should also attract the tape in the same way a battery does. (*b*) The tape and the power cord attract each other. Turning the lamp on or off doesn't make any difference. It does not appear that the power cord has an effect different from any other object.

22.2 (*a*) You should see that the two strips repel each other quite strongly. (*b*) No. Regardless of how you orient the strips—sticky sides facing, nonsticky sides facing, one sticky side facing one nonsticky side—the two always repel. As you bring them closer, they twist to avoid contact.

22.3 Paper-paper, no interaction; tape-paper, interaction; tape-tape, interaction.

22.4 (*a*) The strip should be attracted to your hand and the strip of paper and be repelled by the second charged strip. (*b*) The uncharged strip should remain motionless when you hold your hand close to it. (*c*) You should see no interaction between the uncharged hanging strip and the strip of paper, and there should be an attractive interaction between the two tape strips. This tape-tape interaction is just

like the attractive interaction between a charged tape strip and any other uncharged object.

22.5 The recharged strip should be attracted to your hand and to a strip of paper and repelled by another charged tape strip.

22.6 (*a*) There is an attractive interaction between both types of strips and uncharged objects. (*b*) A third charged strip attracts strip B and repels strip T. (*c*) Yes. Strip T must be charged because it interacts with other charged strips and with uncharged objects. (*d*) Yes. Strip B must be charged because it interacts with other charged strips and with uncharged objects. (*e*) After discharging, neither strip B nor strip T interacts with uncharged objects, and both strips are attracted to a charged strip. (See the box "Troubleshooting B and T strips" if your experimental results do not agree with these.)

Troubleshooting B and T strips

If your experiment with B and T strips doesn't work as expected, check the following:

1. You must pull off the combination in step 2 of Figure 22.5 *very* slowly. (The amount of charge that builds up on the strips is roughly proportional to the speed at which you separate them.) Be sure to remove *all* charge before proceeding.
2. Separating the B and T strips, on the other hand, must be done fairly rapidly. (If you do it too fast, however, so much charge may build up on your strips that it becomes hard to prevent them from being attracted to your hands. If they curl around and touch your hands, you must start over.)
3. Avoid any air currents on the suspended strips.
4. If the humidity of the air is high, the strips may lose their charge rapidly; you may need to repeat the experiment in a drier environment.

22.7 See Figure 22.6.

22.8 (*a*) The comb attracts the uncharged paper strip and repels the (charged) tape strip. (*b*) The comb repels one strip (usually the B strip) and attracts the other. (*c*) The charge on the comb behaves the same way as the charge on the strip it repels (usually B). The charge on the comb is therefore the same type as the charge on the strip it repels.

22.9 If in Checkpoint 22.8 your comb repelled the B strip, your brand of tape produces B strips that carry a negative charge. If the comb attracted the B strip, your brand of tape produces B strips that carry a positive charge.

22.10 (*a*) The red marbles, all carrying a positive charge, exert repulsive forces on one another and so move as far away as possible from one another. (*b*) The blue marbles, all carrying a negative charge, also exert repulsive forces on one another and so move as far as possible from one another. (*c*) The red and blue marbles are attracted to each other and so form red-blue pairs. The (sum of the) charge on any given pair is zero and so does not repel the other pairs. Consequently the pairs do not spread out. (The positive and negative charges do not overlap completely—a red marble and a blue one cannot be at the same place—and so, as we shall see later, some residual interaction is left.) (*d*) Each blue marble becomes part of a red-blue pair; the leftover red marbles repel one another and spread out. (*e*) Because of the surplus of positively charged red marbles, the entire collection carries a positive charge.

22.11 (*a*) Any charge deposited on the metal rod spreads out over the entire conducting system, which in this case is the metal rod plus your body. Most of the charge therefore ends up on you. (*b*) Because rubber is an electrical insulator, any charge on the rod that is not in contact with your hand stays on the rod.

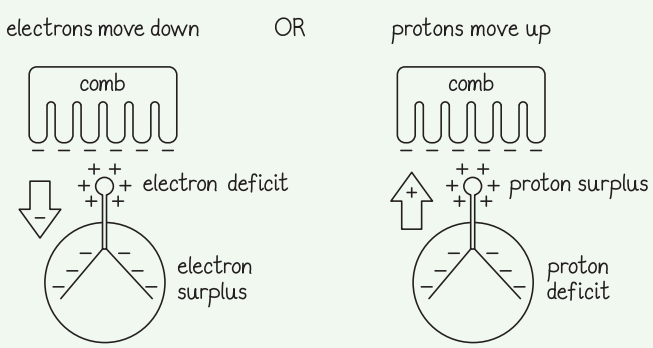
22.12 When you rub together surfaces made of the same material, friction does occur, which tells you that bonds do form between the surfaces. Because the two surfaces are made of the same material, however, there is no preferred direction to transfer charge—the same numbers of electrons are transferred in each direction.

22.13 (*a*) Before the charged rod is brought nearby, the electroscope is neutral, meaning it has as many positive charge carriers as negative charge carriers. No charge is transferred to the electroscope, and so even with the rod nearby, the electroscope still has as many positive charge carriers as negative charge carriers (although they are now separated) and is still neutral. (*b*) The magnitudes must be equal because otherwise the electroscope would not be neutral. (*c*) The negative charge carriers on the rod attract the positive charge carriers on the electroscope ball and repel the negative charge carriers on the leaves. (*d*) The positive charge on the ball is equal in magnitude to the negative charge on the leaves. Therefore, all other things being equal, the magnitudes of the forces are the same. However, the distance between leaves and rod is greater than the distance between ball and rod. Because the electric force decreases with increasing distance, the magnitude of the attractive force exerted on the ball will be *greater* than the magnitude of the repulsive force exerted on the leaves.

22.14 The electron cloud is attracted not only to the external positive charge but also to the positive nucleus. The attractive electric interaction between electron cloud and nucleus prevents the electron cloud from leaving the atom entirely.

22.15 (*a*) Attractive. When the external negatively charged body in Figure 22.21 is replaced by a positively charged body, the atoms in the paper still get polarized. The polarization is now in the other direction, however, so that the negative charge appears on the top in Figure 22.21, meaning the force is again attractive. The vector sum of the forces exerted by the charged object on the piece of paper is still attractive. (*b*) It makes no difference whether positively charged protons move up or negatively charged electrons move down. The effect is the same: a surplus of positive charge on the ball of the electroscope and a deficit of positive charge on its leaves (Figure S22.15). (*c*) No, because the outcome does not depend on the type of charge that moves. The experiment that demonstrates that it is indeed the electrons that flow in a metal—as I asserted in Section 22.3—is beyond the scope of this chapter. We will discuss it in Chapter 27.

Figure S22.15



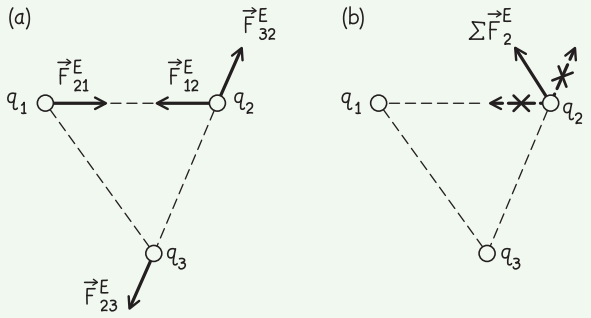
22.16 Greater. For simplicity we treat the spheres as particles so we can apply Coulomb's law. Before the spheres touch, the magnitude of the force exerted by them on each other is $k(q)(3q)/d^2 = 3kq^2/d^2$. When the spheres touch, the charge spreads out evenly over them, so each sphere carries a charge of $(q + 3q)/2 = 2q$. At separation d, therefore, the magnitude of the electric force between them is now $k(2q)(2q)/d^2 = 4kq^2/d^2$, which is greater than it was before.

22.17 Two particles carrying like charges repel each other. If you release them, they accelerate away from each other, gaining kinetic energy. Because the two particles form a closed system, this gain in kinetic energy must be due to a decrease in potential energy associated with the electric interaction. Thus the potential energy of a system consisting of two particles carrying like charges decreases as the separation of the two objects increases. Two particles carrying opposite charges attract each other and so accelerate toward each other. As with two particles carrying like charges, this acceleration means a gain in kinetic energy and a decrease in potential energy. The separation distance is decreasing, however, and we are asked about the change in the system's potential energy when the separation distance is *increasing*. We can reason that because the potential energy decreases as the separation distance between two oppositely charged particles decreases, it must increase as the separation distance increases.

22.18 (*a*) The positive charge carriers repel one another, making their separation greater than the distance between the centers of the spheres. Because Coulomb's law has a $1/r^2$ dependence, the greater separation results in a force of smaller magnitude. (*b*) If the two spheres carry opposite charges, the charge carriers attract one another, and so the distance between them becomes smaller than the distance between the centers of the spheres. Consequently, the magnitude of the attractive force is greater than that obtained from Coulomb's law.

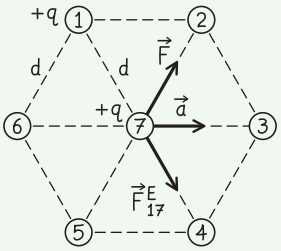
22.19 Because particles 2 and 3 both interact attractively with particle 1, they must carry like charges, and so the interaction between 2 and 3 must be repulsive (Figure S22.19*a*). The vector sum of the forces exerted on 2 is the sum of a leftward horizontal force exerted by particle 1 and a force directed upward and to the right exerted by particle 3; see Figure S22.19*b*.

Figure S22.19



22.20 (*a*) The repulsive force $\vec{F}^E_{17}$ exerted by 1 on 7 points in the direction shown in Figure S22.20. To give 7 an acceleration to the right, the vector sum of the forces exerted on 7 must point to the right. This means that, in addition to $\vec{F}^E_{17}$, another sphere must exert a force $\vec{F}$ on 7 as indicated in the figure. Because the electric force is central (it always acts along the line connecting the two interacting objects), the only two spheres that can exert such a force are 2 and 5. (*b*) The magnitude F must be equal to $\vec{F}^E_{17}$ so that the components in the vertical direction add up to zero. So the magnitude of the charge on sphere 2 or 5 must be the same as that on 1 and 7. To give $\vec{F}$ the direction indicated in the figure, the charge must be negative if it is placed on 2 and positive if it is placed on 5.

Figure S22.20



Chapter 23

23.1 (*a*) You know that the magnitude of the gravitational force exerted by Earth on an object of mass m_o is $F^G_{Eo} = Gm_Em_o/r^2_{Eo}$, where G is the gravitational constant and r_{Eo} is the Earth-object distance (Eq. 13.2). Because G and m_E are constants and r_{Eo} is the same for both objects, the magnitude of F^G_{Eo} is proportional to m_o, and thus $F^G_{E1}/F^G_{E2} = m_1/m_2$. (*b*) Yes, because mass is the only property of the objects in the gravitational force expression. Dividing by m_o removes that one property from the expression: $(Gm_Em_o/r^2_{Eo})/m_o = Gm_E/r^2_{Eo}$.

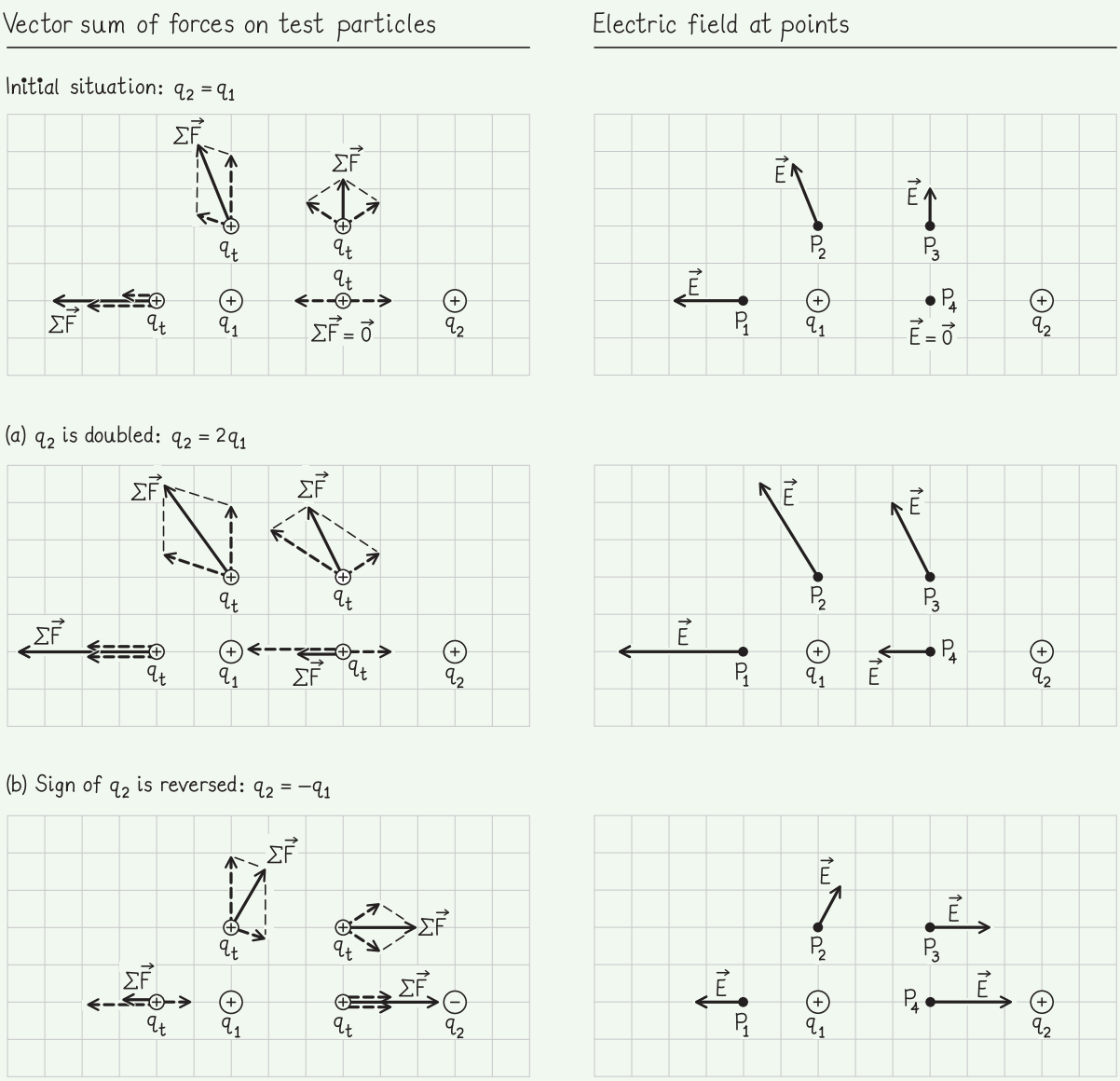
23.2 (*a*) The gravitational field at any location gives the gravitational force exerted on each kilogram of an object at that location. Because each kilogram of the satellite is subject to a force of 2.0 N, the satellite is subject to a force of magnitude $F^G_{Es} = (2.0\ \text{N/kg})(2000\ \text{kg}) = 4000\ \text{N}$. (*b*) $F^G_{Eb} = (2.0\ \text{N/kg})(0.20\ \text{kg}) = 0.40\ \text{N}$.

23.3 (*a*) The magnitudes are the same because the forces form an interaction pair. (*b*) Earth's gravitational field magnitude is greater than that of the ball because the gravitational pull exerted by Earth on objects is much greater than the pull of the ball on other ordinary objects. (*c*) The two forces are of equal magnitude because it is the *product* of the masses that determines the force magnitude: $F^G_{Eb} = Gm_Em_b/r^2_{Eb}$. The magnitude of the gravitational force exerted by a gravitational field on an object is equal to the product of the magnitude of the field at the location of the object and the mass of the object. Thus F^G_{Eb} is equal to the product of the magnitude g of Earth's gravitational field (proportional to m_E and large) and the (small) mass of the ball m_b. F^G_{bE} is equal to the product of the gravitational field of the ball (proportional to m_b and small) and the (large) mass of Earth m_E.

23.4 (*a*) No. $F^E_{pi} = kqq_i^2/d$ (Eq. 22.1) tells you that the magnitude of the force exerted by the particle on an object i is proportional to the charge on the object; because $q_1 \neq q_2$, the force magnitudes cannot be equal and therefore the forces cannot be equal. (*b*) The acceleration of the particle, from $\vec{a}_i = \vec{F}^E_{pi}/m_i$. (*c*) No, because the forces and mass of the particles are different (unless the difference in mass compensates for the difference in force). What is the same for both objects is $F^E_{pi}/q_i = kq/d^2$.

23.5 (*a*) Both forces point toward the particle because objects that carry charges of different types attract. (*b*) The electric field points in the same direction as the force exerted on a positively charged particle, so it points toward the particle. (*c*) Because objects that carry like charges repel, $\vec{F}^E_{p2}$ points away from the particle. The direction of the electric field, however, still points toward the particle. (*d*) The

Figure S23.7



field points away from the particle because now the force exerted by the particle on a positively charged test particle is repulsive. (*e*) The two electric field magnitudes are the same because the magnitude of the force exerted on a test particle is proportional to the magnitude of the source charge and $|+q| = |-q|$.

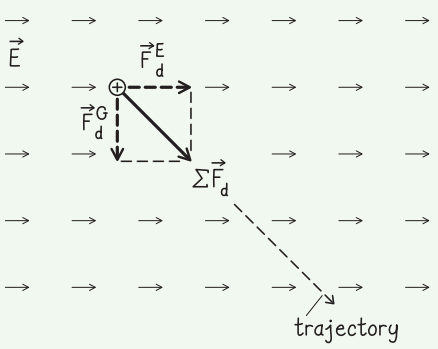
23.6 The magnitude of the force is the product of the electric field magnitude $E = |\vec{E}|$ and the magnitude of the charge q. If q is positive, the direction of the force is the same as that of $\vec{E}$; if q is negative, the force points in the direction opposite the direction of $\vec{E}$.

23.7 (*a*) If q_2 is doubled, the force exerted by particle 2 doubles. Both the direction and the magnitude of $\Sigma\vec{F}^E$ change, and thus both the direction and the magnitude of $\vec{E}$ change (Figure S23.7*a*). (*b*) See Figure S23.7*b*.

23.8 (*a*) In the same direction as the electric field at that point, that is, to the left. (*b*) The magnitude becomes three times as great; the direction is the same. (*c*) Yes, because the force exerted by object 3 changes. This change affects both the direction and the magnitude of $\Sigma\vec{F}^E$. (*d*) Opposite the direction of the electric field at that point, that is, down the page.

23.9 The droplet is subject to a downward force $\vec{F}_d^G$ and a horizontal force $\vec{F}_d^E$. The vector sum of these forces is downward at an angle, as illustrated in Figure S23.9. Both forces are constant in magnitude, so their vector sum is also constant. Consequently the droplet has a constant acceleration along the direction of the vector sum. Because the droplet begins at rest, its trajectory is a straight line along this direction.

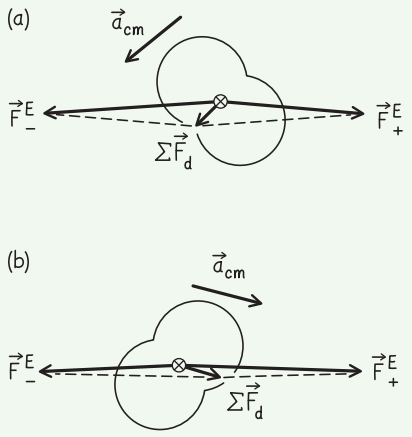
Figure S23.9



23.10 (*a*) The torque causes the molecule to rotate in the counterclockwise direction. (*b*) No. When the dipole axis that passes through the centers of the two particles is parallel to the electric field, the torque is zero because the two forces lie along the dipole axis and the lever arm of the torque is zero. When the axis is perpendicular to the electric field, the torque is a maximum.

23.11 (*a*) See Figure S23.11*a*. The vector sum of the forces—and therefore the center-of-mass acceleration—point to the left and down. (The force exerted on the negative end is greater than that on the positive end because the electric field is stronger there and both forces have a downward component.) (*b*) See Figure S23.11*b*. The acceleration now points to the right and down, and so the dipole begins by moving in that direction (*away* from the source of the electric field). The orientation of the dipole changes, however, because the forces cause a torque. As we saw in part *a*, as soon as the negative end of the dipole is more toward the left than the positive end, the dipole accelerates toward the left.

Figure S23.11



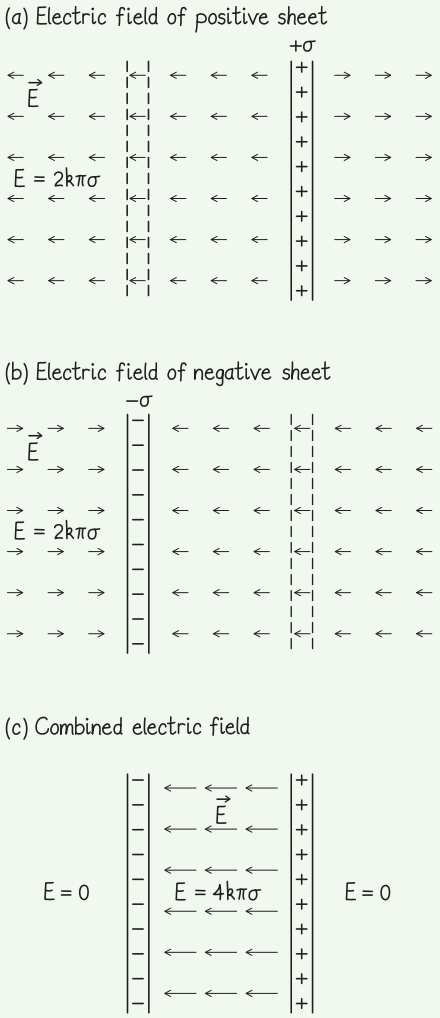
23.12 The magnitude of the electric field is $E = \sqrt{E_x^2 + E_y^2}$. Substituting the values we obtained for E_x and E_y in Example 23.3, we get $E = 1.9 \times 10^4$ N/C. The magnitude of the electric force is $F_e^E = eE = (1.6 \times 10^{-19}\ \text{C})(1.9 \times 10^4\ \text{N/C}) = 3.0 \times 10^{-15}$ N. This force causes an acceleration of magnitude $a_e = F_e^E/m_e = (3.0 \times 10^{-15}\ \text{N})/(9.1 \times 10^{-31}\ \text{kg}) = 3.3 \times 10^{15}\ \text{m/s}^2$. As this example shows, the accumulation of just tens of microcoulombs causes a phenomenally large acceleration of the electron even when it is meters away.

23.13 Equations 23.10 and 23.13 show that the magnitude of the electric field of a dipole is proportional to the dipole moment magnitude p. So the dipole that creates the stronger field (dipole B) has the greater dipole moment. The magnitude of the dipole moment is the product of dipole separation d and dipole charge q_p, which means the greater dipole moment of B can be due either to greater d or to greater q_p. Because we cannot separate these two effects, we cannot say which dipole has the greater dipole charge.

23.14 To determine the electric field created by two parallel charged sheets, you must use the superposition principle. Figure S23.14*a* and *b* shows the electric fields of a positive and a negative sheet. In Figure S23.14*c* these two fields are added. (*a*) Between the sheets the electric fields are in the same direction, and so they reinforce, giving $E = 4k\pi\sigma$ between the sheets. (*b*) Outside the sheets, the electric fields are in opposite directions and therefore add to zero.

23.15 As you saw in Section 22.5, a direct consequence of the $1/r^2$ dependence in Coulomb's law is that the force exerted on a charged particle inside a hollow uniformly charged sphere is zero. This means that the electric field inside such a sphere is zero. For the electric field at some point P inside a uniformly charged sphere of radius R (Figure S23.15), the part of the sphere farther from the center than

Figure S23.14

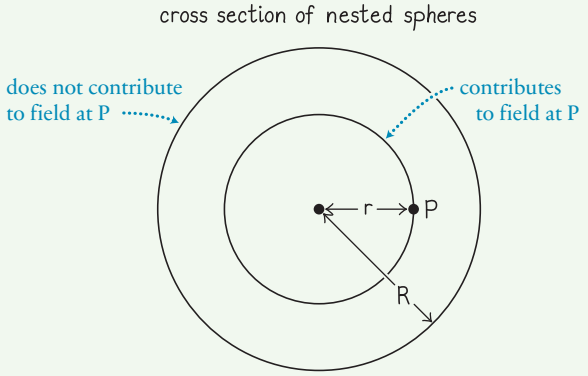


P does not contribute to the electric field at P. If the sphere carries a uniformly distributed charge q, the amount of charge on the part of the sphere closer to the center than P is $(r^3/R^3)q$, and the electric field due to this charge is

$$E_r = k\frac{r^3}{R^3}\frac{q}{r^2} = k\frac{q}{R^3}r.$$

In other words, the electric field *increases* in proportion to the distance r from the center. At the center, it is exactly zero.

Figure S23.15

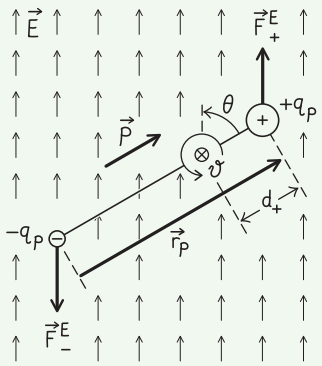


23.16 Yes. To see why, suppose the center of mass is off-center, as shown in Figure S23.16; everything else, including the magnitude of the charges and the dipole moment, is as before. Let the distance between the positive end and the center of mass be d_+. The force exerted on the positive end then causes a counterclockwise torque of magnitude $\tau_+ = r_\perp F_+^E = (d_+ \sin\theta)(q_pE)$, and the force exerted on the negative end causes a counterclockwise torque of magnitude $\tau_- = r_\perp F_-^E = [(d - d_+)\sin\theta](q_pE)$, where $d = |\vec{r}_p|$. Adding the torques yields

$$\begin{aligned}\Sigma\tau_\vartheta &= (d_+ \sin\theta)(q_pE) + [(d - d_+)\sin\theta](q_pE)\\ &= d\sin\theta(q_pE) = (q_pd)E\sin\theta = pE\sin\theta,\end{aligned}$$

identical to the result in Eq. 23.20.

Figure S23.16



23.17 (*a*) Doubles the force (Eq. 23.23). (*b*) Doubles p and thus doubles the force. (*c*) Doubles p and thus doubles the force. (*d*) Reduces the force by a factor of $\left(\frac{1}{2}\right)^3$ because of the $1/y^3$ dependence in Eq. 23.23.

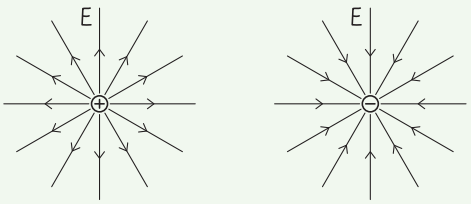
23.18 The dipole moment $\vec{p}$ points in the same direction as the electric field that induces it, so α in Eq. 23.24 must be positive.

23.19 (*a*) Doubles the force (Eq. 22.1). (*b*) Quadruples the force because of the q^2 dependence in Eq. 23.26. (*c*) Doubling q_A in part *b* doubles the electric field created by *A* at the position of the induced dipole. This doubles the magnitude of p_{ind} (the stronger electric field increases the charge separation). This doubled dipole moment then interacts with a doubly strong electric field, increasing the force by a factor of 4. (*d*) No, because the polarization induced by the electric field is always along the electric field (Eq. 23.24). In other words, the induced dipole moment is parallel to the electric field, and the vector product of $\vec{p}$ and $\vec{E}$ is then zero.

Chapter 24

24.1 See Figure S24.1.

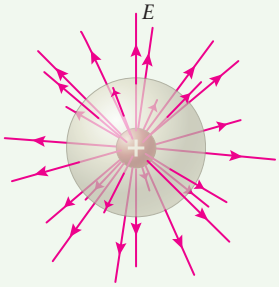
Figure S24.1



24.2 (*a*) No, because at the point of intersection the direction of the electric field would not be unique—it cannot be tangent to *both* intersecting lines at the same time. (*b*) No. Although two field lines that touch have the same tangent at the point where they touch, the two field lines would have different tangents on either side of that point, and so the direction of the electric field would again not be unique.

24.3 (*a*) As illustrated in Figure S24.3, the same 26 field lines pass through the surface of the hollow sphere. (*b*) The surface area of a sphere of radius R is $4\pi R^2$, so the number of field lines per unit area is $(26)/(4\pi R^2)$. (*c*) Again 26. (*d*) The number of field lines per unit area on the second sphere is $(26)/[4\pi(2R)^2] = (26)/(16\pi R^2)$, which is reduced by a factor of 4 from that for the sphere of radius R. (*e*) The electric field decreases as $1/r^2$, so doubling the distance reduces the electric field by a factor of 4.

Figure S24.3



24.4 (*a*) No field lines cross the surface when it is parallel to the field lines. (*b*) If 16 field lines pass through an area of 1 m^2, then 8 field lines pass through an area of 0.5 m^2. (*c*) The number of field lines per unit area is $(8)/(0.5\ m^2) = 16\ m^{-2}$, or 16 field lines per square meter. (*d*) $(16)/(1\ m^2) = 16\ m^{-2}$, as above. The field line density is the same because the field is uniform.

24.5 (*a*) It remains the same—each field line still passes through the surface of the sphere. (*b*) Because neither the number of field lines nor the surface area of the sphere changes, the average number of field line crossings per unit surface area remains the same. (*c*) The electric field strength increases as the distance to the charged object decreases, so moving the sphere off-center increases the electric field strength on one side and decreases it on the opposite side. (*d*) No. The answer to part *b* gives the *average* field line density. Moving the sphere off-center increases the field line density on one side and decreases it on the other, so the average field line density can remain the same.

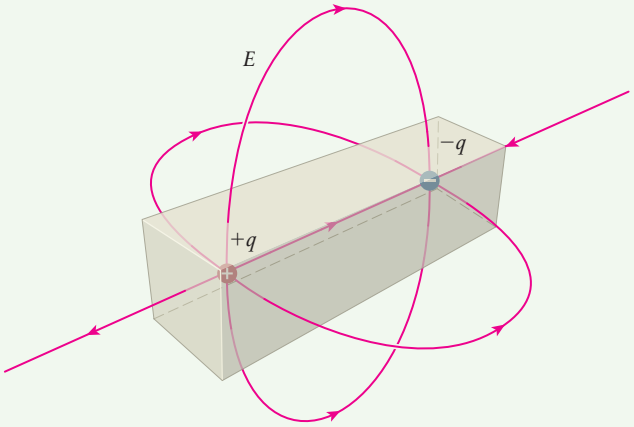
24.6 (*a*) Sixteen field lines emanate from the object, so 16 field lines pass through the surface of the sphere. (*b*) Eight field lines emanate from each of the objects, so 16 field lines pass through the surface of the sphere. (*c*) The amount of charge enclosed by the sphere is (20 field lines)/(8/q field lines per unit charge) = 2.5q.

24.7 (*a*) Each field line that reenters the donut must also exit, contributing a value of $(+1) + (-1) = 0$ to the field line flux. Thus, regardless of how many field lines reenter the donut, the field line flux remains 6. (*b*) No. Regardless of the shape of the surface, each field line contributes no more and no less than +1 to the field line flux, so the field line flux is always equal to 6.

24.8 (*a*) The field line flux is zero because each field line that enters the sphere also exits the sphere, and so does not contribute to the field line flux. (*b*) No. Moving the particle changes the number of field lines that enter the sphere, but regardless of how many field lines enter the sphere, each of them must also leave the sphere, so the field line flux remains zero.

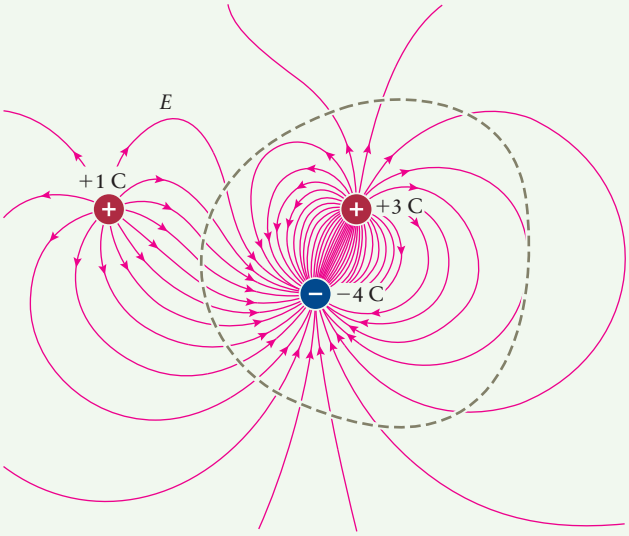
24.9 Five field lines emanating from the positively charged particle contribute a flux of +5 (the field line that points from the positive end to the negative end doesn't pass through the surface; see Figure S24.9). Five field lines terminating on the negatively charged particle contribute a flux of −5. This gives a flux of zero, which makes sense because the amount of charge enclosed by the box is $+q + (-q) = 0$. [Notice that this answer remains valid even if we make the box much larger so that all the curved field lines fit inside it. Then only two field lines pass through the box—one emanating from the positively charged particle, the other terminating on the negatively charged one. The flux is then still zero: $(+1) + (-1) = 0$.]

Figure S24.9



24.10 (*a*) Fifteen field lines go into the surface and three come out of it. The field line flux is thus $(-15) + (+3) = -12$. Twelve field lines emanate from the object carrying a charge of +1 C, so the charge inside the hidden region must be −1 C. Figure S24.10 shows the entire field line diagram; as you can see, the charge enclosed by the dashed line is indeed $(+3\text{ C}) + (-4\text{ C}) = -1\text{ C}$. (*b*) Going along the perimeter of the illustration, we note that four field lines leave the edge of the illustration and four enter it. The field line flux is thus $(+4) + (-4) = 0$. This makes sense because the charge enclosed within the diagram is $(+1\text{ C}) + (-1\text{ C}) = 0$.

Figure S24.10



24.11 If there were a flux into the shell from the left and out of the shell on the right, then the situation would no longer have the required symmetry—if you were to rotate the sphere 180° about the vertical axis, the field line flux would be reversed and therefore the situation would be different. Because the situation is spherically symmetrical, this case is not possible.

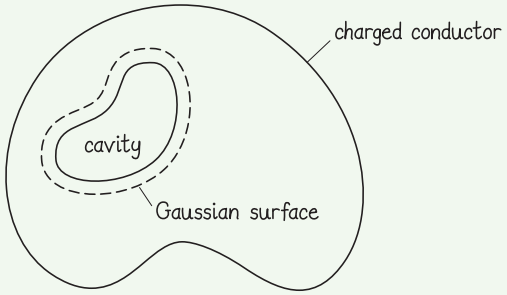
24.12 (*a*) It stays the same because the field line density is the same everywhere on that surface. (*b*) It decreases because the field line spacing increases. (*c*) Zero. The field lines are all perpendicular to the wire, so no field lines pass through those surfaces.

24.13 (*a*) It stays the same because the field line density is the same everywhere along a plane parallel to the charged sheet. (*b*) It stays the same because the field line spacing doesn't change. (*c*) Positive, because the field lines cross from inside the surface to outside. (*d*) They are the same in magnitude and in sign. (*e*) Zero. The field lines are all perpendicular to the sheet, so no field lines pass through the curved surface.

24.14 (*a*) The field line flux is zero because the field is zero everywhere inside the conducting object. (*b*) According to the relationship between field line flux and enclosed charge derived in Section 24.3, the charge enclosed by the surface must be zero.

24.15 No. Imagine a Gaussian surface around the cavity, just inside the conducting object (Figure S24.15). Because $E = 0$ everywhere inside the object, the field line flux through this Gaussian surface is zero and so the charge enclosed by it must also be zero. So, the surplus charge on the object cannot reside on an inner surface.

Figure S24.15

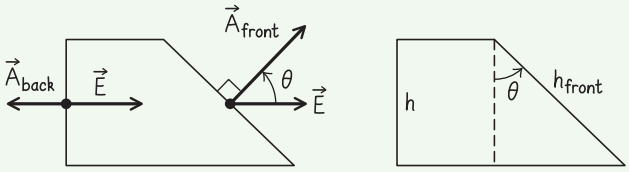


24.16 No. Field lines emanate from the positively charged particle and terminate on the negative charge carriers that line the cavity wall. Thus, "$E = 0$ everywhere inside a conducting object" means everywhere inside the bulk of the object. Cavities don't count!

24.17 (*a*) (*i*) Increases (more field lines intercepted); (*ii*) increases (greater field line density so more field lines intercepted); (*iii*) increases (more field lines intercepted). (*b*) (*i*) Increases (greater A yields greater Φ_E, by Eq. 24.1); (*ii*) increases (greater $|\vec{E}|$ yields greater Φ_E); (*iii*) increases (greater slope means smaller θ, which yields a greater $\cos\theta$).

24.18 (*a*) The area of the back surface is 1.0 m^2, so the magnitude of the corresponding area vector $\vec{A}_{\text{back}}$ is 1.0 m^2. From Figure S24.18 we see that $h/h_{\text{front}} = \cos\theta = \cos 30° = 0.87$, so the area of the front surface is larger by a factor of $h_{\text{front}}/h = 1.2$. The magnitude of the area vector $\vec{A}_{\text{front}}$ is thus 1.2 m^2. (*b*) We use Eq. 24.2 to calculate the electric fluxes. For the back surface we get $\Phi_E = EA_{\text{back}} \cos(180°) = (1.0\text{ N/C})(1.0\text{ m}^2)(-1) = -1.0\text{ N}\cdot\text{m}^2/\text{C}$; for the front surface we have $\Phi_E = EA_{\text{front}} \cos\theta = (1.0\text{ N/C})(1.2\text{ m}^2)(0.87) = 1.0\text{ N}\cdot\text{m}^2/\text{C}$.

Figure S24.18

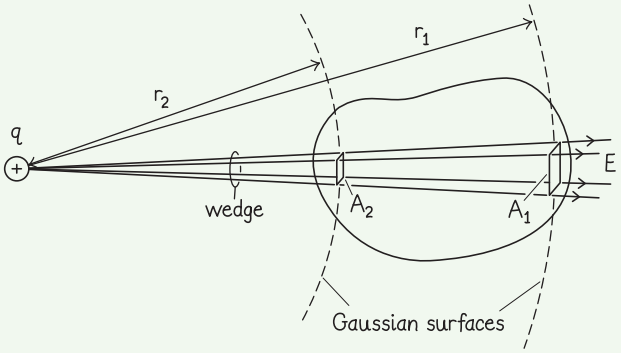


24.19 (*a*) At a distance r from a particle carrying a charge $+q$, the magnitude of the electric field is kq/r^2 (see Eq. 23.4). (*b*) Because the electric field is perpendicular to the sphere at all points and has the same constant magnitude, the electric flux is given by the product of the electric field and the area of the sphere: $\Phi_E = EA = E(4\pi r^2) = (kq/r^2)(4\pi r^2) = 4\pi kq$. (*c*) The enclosed charge is $+q$, so $\Phi_E = 4\pi kq_{enc}$. (*d*) No. If we double the radius r, the field decreases by a factor of 4 but the area of the sphere increases by a factor of 4. Thus the electric flux and its relationship to q_{enc} remain the same.

24.20 (*a*) With the modified Coulomb's law, the magnitude of the electric field at the surface of the sphere would be $kq/R^{2.00001}$. The electric flux thus becomes $\Phi_E = EA = E(4\pi R^2) = (kq/R^{2.00001})(4\pi R^2) = 4\pi kqR^{-0.00001}$. (*b*) $4\pi kqR^{-0.00001} = q_{enc}/\epsilon_0 = 4\pi kq_{enc}$ or $qR^{-0.00001} = q_{enc}$, which is not an equality. The cancellation of the radius R happens only when the dependence on r in Coulomb's law is *exactly* $1/r^2$.

24.21 Figure S24.21 shows how a wedge from q intersects the surface. The electric flux through intersection A_1 is equal to the electric flux through the intersection of the wedge with a sphere of radius r_1 centered on q. Likewise, the electric flux through A_2 is equal to the electric flux through the intersection of the wedge with a sphere of radius r_2 centered on q. As we saw in Checkpoint 24.19, the magnitudes of the electric flux through a sphere are independent of the radius. Because the intersections of the wedge with the two concentric spheres represent a fixed fraction of the total surface area of the spheres, the electric fluxes through these two intersections must also be the same. Consequently, the magnitudes of the electric fluxes through A_1 and A_2 are the same. Because the electric field points out of the sphere at A_1 and into the sphere at A_2, however, the algebraic signs of the electric fluxes are opposite: $\Phi_{E1} = -\Phi_{E2}$, so $\Phi_{E1} + \Phi_{E2} = 0$. Because this holds for *any* wedge, we learn that the electric flux through the surface due to q is zero when q is outside the surface.

Figure S24.21



24.22 The electric field is also given by Eq. 24.15 because none of the arguments leading up to that result change: The field is still spherically symmetrical, and the enclosed charge is still $+q$.

24.23 No, because such a charge configuration does not have sufficient symmetry to make the calculation practical. A finite rod still has *rotational* symmetry (you can rotate it about its axis without changing the configuration), but it cannot be translated about its axis without changing the configuration. For a finite rod the field changes as you move up and down parallel to the y axis in Figure 24.30. With only rotational symmetry, we cannot easily find a closed surface such that for the separate regions of the surface either the magnitude of the field is constant or the electric flux is zero.

24.24 (*a*) The charge spreads out over the outside surface area of the plate. If we ignore the edges of the plate, $\frac{1}{2}q$ spreads out over each side of surface area A, and so the surface charge density (charge per unit surface area) is $\frac{1}{2}q/A$ on each side. The magnitude of the field is given by Eq. 24.17: $E = \sigma/\epsilon_0 = (\frac{1}{2}q/A)/\epsilon_0 = q/(2\epsilon_0 A)$. (*b*) The surface charge density is q/A; the magnitude of the electric field is given by the solution of Exercise 24.8: $E = \sigma/(2\epsilon_0) = (q/A)/(2\epsilon_0) = q/(2\epsilon_0 A)$, which is the same result as in part *a*. The point to remember therefore is that the difference between the solutions of Exercise 24.8 and Eq. 24.17 arises solely from the difference in the surface charge that should be used in each of them.

Chapter 25

25.1 (*a*) The kinetic energies are the same because the two particles are subject to the same force and the force displacements are the same, so the work done on them is the same. (*b*) Particle 2 has the greater momentum. As you may recall from Chapter 8, the change in momentum of an object is given by the product of the force exerted on it and the time interval during which the force is exerted. Because particle 2 has greater mass, its acceleration is smaller, and so it takes longer to fall the same distance as particle 1. (*c*) Both have the same momentum (because the product of force and time interval is the same for both), but particle 1 has greater kinetic energy (because particle 1 has smaller mass, its acceleration is greater and so it undergoes a greater displacement than particle 2; therefore the work done on it is greater and it gains more kinetic energy). (*d*) The electric force exerted on each particle is equal to the product of the electric field and the charge on the particle. Thus, their accelerations are given by

$$\vec{a}_1 = \vec{F}_1^E/m_1 = q_1\vec{E}/m_1$$
$$\vec{a}_2 = \vec{F}_2^E/m_2 = q_2\vec{E}/m_2.$$

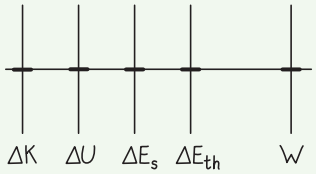
Equating these two accelerations yields $q_1/m_1 = q_2/m_2$ or, rearranging terms, $q_1/q_2 = m_1/m_2$. In other words, if we make the ratio of the charges on the particles equal to the ratio of their masses, their accelerations are the same.

25.2 (*a*) When the dipole is vertical, the electric potential energy has reached a minimum because the separation between the positive end of the dipole and the charged object is at a maximum and the distance between the negative end and the object is at a minimum. As the dipole rotates past the vertical, the electric potential energy increases. (*b*) When the dipole is vertical, the torque due to the electric field is zero. Because of its (rotational) kinetic energy, however, the dipole continues to rotate beyond this point and the torque due to the electric field reverses direction. The rotation of the dipole therefore slows down and (rotational) kinetic energy is converted back to electric potential energy. The dipole comes to rest when all of its kinetic energy has been converted—this happens when it is again horizontal (but now with the positive end to the right). The motion then reverses and the dipole continues to oscillate back and forth. (*c*) The same principles apply, but the oscillation would take place over different angles. In other words, if the dipole started at an angle of 45° to the vertical, then the oscillation would occur between $\pm 45°$ instead of $\pm 90°$ as in part *b*.

25.3 (*a*) The work done by the electric field on the particle is the product of the x components of the force exerted on it and the force displacement. The x component of the particle's displacement is $x_B - x_A$ and the x component of the force exerted on it is qE_x, so $W_{Ep}(A \to B) = qE_x(x_B - x_A)$. (*b*) As the particle is moved back to its initial position, the force displacement is reversed, so the work done by the electric field on the particle is the negative of the amount done in part *a*: $W_{Ep}(B \to A) = -qE_x(x_B - x_A)$. (*c*) There is no change in

the particle's kinetic energy and the particle possesses no internal energy, so its energy does not change as the agent moves it back. Therefore the agent must do an amount of work on the particle that is the negative of the work done by the electric field on it calculated in part *b*: $W_{ap}(B \to A) = -W_{Ep}(B \to A) = +qE_x(x_B - x_A)$. (*d*) The sum of the work done by the agent and by the electric field on the particle is zero. (*e*) See Figure S25.3. (There are no changes in the particle's energies and so the total work done on the particle is zero.)

Figure S25.3



25.4 (*a*) The electrostatic work done on a particle as it moves along the gray path from A to C is equal to the electrostatic work done on it from A to B: $W_{Ep}(\text{gray path}) = W_{Ep}(A \to B)$. The work done from B to C is zero, as is that from C to B: $W_{Ep}(B \to C) = W_{Ep}(C \to B) = 0$. Thus, the electrostatic work done on the particle from C to B to A is just that done from B to A: $W_{Ep}(CBA) = W_{Ep}(B \to A)$. Because the electrostatic work done from B to A is the negative of that done from A to B (which is equal to W), the electrostatic work done from C to B to A must be $-W$. (*b*) Adding the result we found in part *a* to the electrostatic work done along the gray path, we calculate for the work along the closed path from A to C to B and back to A $W + (-W) = 0$.

25.5 (*a*) No. See Checkpoint 25.4. Although the electrostatic work done along the entire path is zero, that along segment AC is nonzero. (*b*) No. The electric field varies in magnitude (and therefore the electric force varies too) along the path AB. We must use integral calculus to calculate the work done by a variable force (see Section 9.7). (*c*) (*i*) If the charge on the particle is doubled, then the electric force exerted on the particle doubles and so the electrostatic work done on it doubles. (*ii*) The electric force does not depend on the mass m of the particle, so the electrostatic work done on it, too, is independent of m.

25.6 (*a*) Negative, because the electrostatic work done on a positively charged particle moving along any path from A to C is positive and the potential difference is the negative of the electrostatic work done per unit charge. (*b*) Zero, because the electrostatic work done on the particle is zero along any path from C to B. (*c*) Positive, because the electrostatic work done on the particle is negative (along the straight path from B to A the force and force displacement are in opposite directions). (*d*) The electrostatic work done on the particle is positive because the electric force and the force displacement are in the same direction. The potential difference is therefore negative. (*e*) The electrostatic work done on the particle is equal to the change in kinetic energy: $W_{Ep} = \Delta K$. This is the electrostatic work done on a particle carrying a charge q, and so the electrostatic work done per unit charge is $\Delta K/q$. The potential difference is the negative of this quantity: $-\Delta K/q$.

25.7 Yes. The surface of any sphere that is centered on the particle constitutes an equipotential surface. (The electric field is always perpendicular to such a surface, and so the electrostatic work done on a particle as it is moved along any path—regardless of its shape—that lies on such a surface is always zero.)

25.8 Zero. As we have seen in Section 24.5, the electric field inside a conducting object that is in electrostatic equilibrium is zero. The accumulation of positive and negative charge carriers on opposite ends of the sphere occurs precisely to cancel the effect of the external field of the charged rod anywhere inside the sphere: The electric field caused by the polarization of charge on the sphere exactly cancels the electric field of the rod. Therefore, because $\vec{E} = 0$ inside the conducting object, the electrostatic work done on a charged particle inside the object is zero, too. Consequently the potential difference between two points on or inside the object is zero.

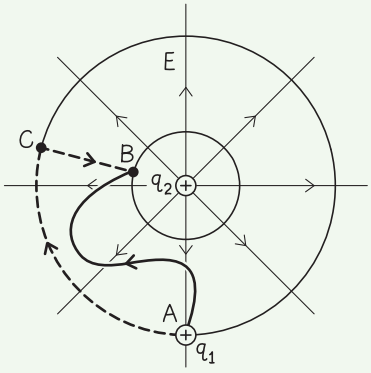
25.9 (*a*) Along path CB, $r_i > r_f$, so the left side of Eq. 25.5 is negative. (*b*) Positive work must be done to push a positively charged particle toward another positively charged particle because the force the external agent doing the pushing must exert and the force displacement are in the same direction. Because the kinetic energy of particle 2 doesn't change, the total work done on 2 (the sum of the electrostatic work done by particle 1 on 2 and the work done by the external agent on 2) must be zero: $W_2 = W_{12} + W_{a2} = 0$. So if the electrostatic work done by particle 1 on 2 is negative ($W_{12} < 0$), then the work done by the external agent on 2 must be positive: $W_{a2} > 0$. (*c*) It we take both particles as our system, there is no longer any electrostatic work done on the system (the electric interaction is now internal). Consequently the work done by the external agent in moving 2 changes the electric potential energy of the two-particle system: $\Delta U^E = W_{a2}$. Using our answer to part *b*, we have $W_{a2} = -W_{12}(C \to B)$ or, using Eq. 25.4,

$$\Delta U^E = -W_{12}(C \to B) = -\frac{q_1 q_2}{4\pi\varepsilon_0}\left[\frac{1}{r_C} - \frac{1}{r_B}\right].$$

Because $r_C > r_B$, this change in energy is positive.

25.10 Yes. Figure S25.10 shows particle 1 moving instead of particle 2 (compare with Figure 25.15). We follow the same derivation as in Eq. 25.2, substituting $\vec{F}_{21}^E$ for $\vec{F}_{12}^E$. Because the magnitudes of these two forces are the same, the right-hand side of Eq. 25.2 remains the same. The initial and final separations are also still the same and so we obtain the same end result (Eq. 25.4). Alternatively, you can look at Eq. 25.5 and note that the expression is symmetrical in q_1 and q_2 and that r_{12i} and r_{12f} are independent of which of the two charged particles is moved.

Figure S25.10



25.11 Yes. The electrostatic work done on 3 as it is brought in first is zero: $W_3 = 0$; the electrostatic work done on 1 as it is brought in next is (compare with Eq. 25.9)

$$W_{31} = -\frac{q_3 q_1}{4\pi\varepsilon_0}\frac{1}{r_{31}}.$$

Likewise, the electrostatic work done on 2 when bringing in 2 is (compare with Eq. 25.12)

$$W_{12} + W_{32} = -\frac{q_1 q_2}{4\pi\varepsilon_0}\frac{1}{r_{12}} - \frac{q_3 q_2}{4\pi\varepsilon_0}\frac{1}{r_{32}}.$$

The total electrostatic work $W = W_{31} + W_{12} + W_{32}$ done on the system is still the same.

25.12 (*a*) Positive. The potential difference between A and B is $V_B - V_A$, so A is the initial point and B is the final point: $r_i = r_A$ and $r_f = r_B$. Because $r_A > r_B$, the right side of Eq. 25.20 is positive. (*b*) Along the straight path from A to B the angle between the electric force and the force displacement is between 90° and 180°, so the electrostatic work done on the particle is negative. This is consistent with the answer to part *a*; see Eq. 25.15.

25.13 The electric field of a charged particle is given by Eq. 23.4, $\vec{E}_s(P) = (kq_s/r_{sP}^2)\hat{r}_{sP}$, so Eq. 25.25 becomes

$$V_{AB} = -\int_A^B \vec{E}\cdot d\vec{\ell} = -\frac{q}{4\pi\varepsilon_0}\int_A^B \frac{\hat{r}}{r^2}\cdot d\vec{\ell}.$$

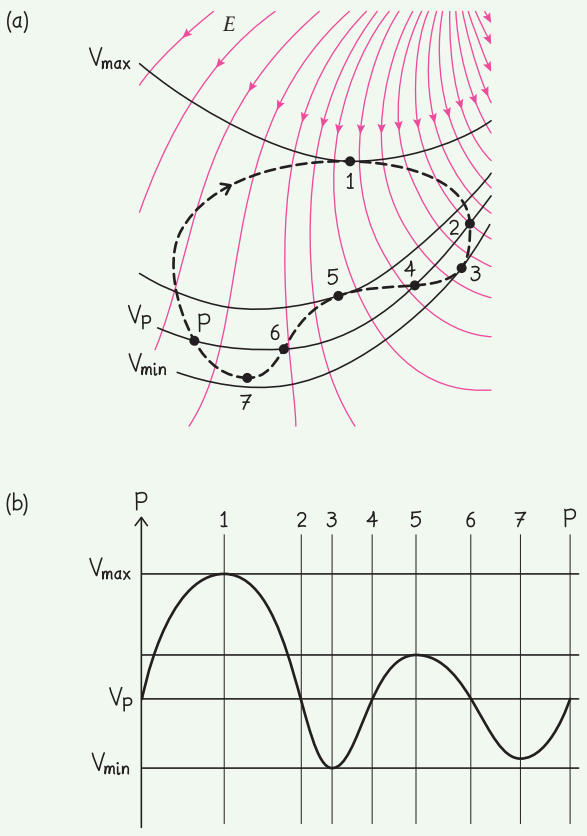
The scalar product on the right-hand side is zero along a circular arc because $\hat{r}$ and $d\vec{\ell}$ are perpendicular. Along a radial line, $\hat{r}$ and $d\vec{\ell}$ are parallel, so $\hat{r}\cdot d\vec{\ell} = dr$. Therefore

$$V_{AB} = -\frac{q}{4\pi\varepsilon_0}\int_{r_A}^{r_B}\frac{dr}{r^2} = \frac{q}{4\pi\varepsilon_0}\left[\frac{1}{r}\right]_{r_A}^{r_B}$$
$$= \frac{q}{4\pi\varepsilon_0}\left[\frac{1}{r_B}-\frac{1}{r_A}\right],$$

which is the same result we obtained in Eq. 25.20.

25.14 We begin by sketching some equipotentials for the electric field pattern shown in Figure 25.22. Because equipotentials are always perpendicular to field lines, we can sketch equipotential lines by always drawing them perpendicular to the field lines, as in Figure S25.14*a*. Based on the equipotential lines I have drawn, I choose a set of representative points to evaluate how the potential varies. Moving clockwise around the path from P, I note that the displacement is upward, whereas the direction of the electric field (and therefore the direction of the electric force exerted on a positively charged test particle) is downward. Thus, the electrostatic work done on the particle is negative and the potential difference between a point above the equipotential passing through P and P is positive. In other words, the potential *increases*. At point 1 we reach maximum potential; moving beyond 1, the potential decreases again. At point 2 we cross the equipotential through point P again (the value of the potential is again V_P). The potential continues to decrease until point 3, where it reaches a minimum value. The potential then increases again until a local maximum at point 5 is reached. As we return to P, the potential reaches another local minimum at 7 before increasing again to the initial value. See Figure S25.14*b*.

Figure S25.14



25.15 If the rod and the disk are positively charged and the zero of potential is at infinity, then both potentials should be *positive* (the electrostatic work done on a positively charged particle is negative as we bring it in from infinity). In the answer to Example 25.6, the factor in the logarithm is greater than 1 because $\sqrt{\ell^2+d^2}$ is greater than d, so the numerator is always greater than the denominator. The logarithm is thus positive, yielding a positive V for positive q. Likewise, in the answer to Example 25.7, the factor in parentheses is always positive because $\sqrt{z^2+R^2}$ is greater than $|z|$ so, for positive σ, the potential is positive.

25.16 In Example 25.5, we found for the potential between the plates, $V(a) = E(d-a)$, where a is the distance from the positive plate. Rewriting this result as a function of x, we have $V(x) = E(d-x)$. Because $V(x)$ is not a function of y and z, the partial derivatives of V with respect to y and z are zero. Thus $E_y = E_z = 0$. Therefore, the electric field must be in the positive x direction, and the component in that direction is

$$E_x = -\frac{\partial V}{\partial x} = -\frac{\partial}{\partial x}[E(d-x)]$$
$$= -\frac{\partial}{\partial x}(Ed) + \frac{\partial}{\partial x}(Ex) = 0 + E = E.$$

The electric field is in the x direction and of magnitude E, $\vec{E} = E\hat{\imath}$, in agreement with the situation shown in Figure 25.20.

25.17 The potential we obtained in Example 25.7 is a function of z only, so the x and y components of the electric field are zero. The z component is given by

$$E_z = -\frac{\partial V}{\partial z} = -\frac{\sigma}{2\varepsilon_0}\frac{\partial}{\partial z}(\sqrt{z^2+R^2}-|z|)$$
$$= -\frac{\sigma}{2\varepsilon_0}\left(\frac{z}{\sqrt{z^2+R^2}}-\frac{z}{|z|}\right).$$

Substituting $k = 1/(4\pi\varepsilon_0)$, we get, for $z > 0$

$$E_z = 2k\pi\sigma\left[1-\frac{z}{\sqrt{z^2+R^2}}\right]$$

which is the same result we obtained by direct integration in Example 23.6 (see Eq. 2). Note how much easier and shorter the derivation of the electric field via the potential is compared to the direct integration of Section 23.7.

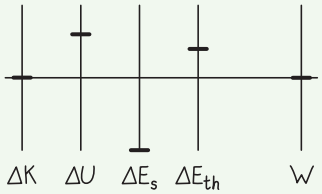
Chapter 26

26.1 (*i*) The force exerted on a unit positive charge at P is the vector sum of the electric forces due to all the charge carriers on the rod and the fur exerted on the unit charge. If the distribution of charge on the rod and the fur is the same as before, then doubling the charge doubles the magnitude of the individual electric forces without changing their direction. Consequently the direction of the electric field should be the same, but its magnitude should be twice as great. (*ii*) Because the electric field increases by a factor of 2, the potential difference (which is equal to the negative of the line integral of the electric field along any path between the two fixed points) increases by a factor of 2 as well. (*iii*) The potential energy stored in the system increases *by more than a factor of 2* because as charge is transferred from the fur to the rod, the charges already on the rod and the fur make the transfer more difficult because their field opposes the transfer. Thus, the transfer of each additional amount of charge requires more energy than the previous amount of charge. Put differently, the second half of the charge is transferred against a much greater potential difference than the first half.

26.2 The charged fur and rod attract each other, so you must do positive work on the rod-fur system to increase the separation of the rod and the fur. This work increases the electric potential energy of the rod-fur system.

26.3 See Figure S26.3. No external agent does any work on the system. The person, now part of the system, provides the energy from source energy.

Figure S26.3



26.4 (*a*) When the wires are disconnected, any charge already on the plates remains there, so the potential difference between the plates remains the same. (*b*) The charging of the capacitor continues until the potential difference across the capacitor is equal to that across the battery. Thus, if the battery has a greater potential difference, then the potential difference between the plates is greater too. This greater potential difference requires more charge on each plate.

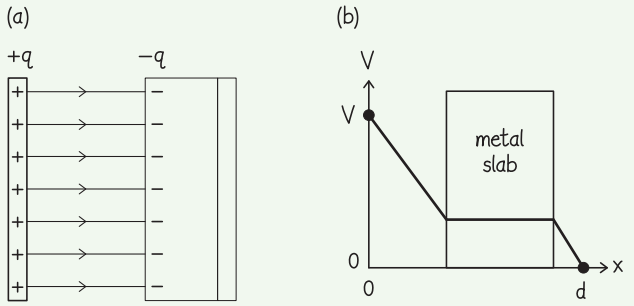
26.5 (*a*) Connecting the capacitors to a 9-V battery means that the potential difference across each capacitor is 9 V. As we saw in Section 25.5, the potential difference across a parallel-plate capacitor is the product of the electric field E and the plate separation d (see Example 25.5). Therefore, the capacitor with the smaller plate separation has the greater electric field. Because the electric field of a charged plate is proportional to its surface charge density (see Exercise 24.8), the plates of the capacitor with the smaller plate separation carry the greater charge. (*b*) The size and shape of the plates don't enter into the expression for the electric field, so the smaller plates must have the same surface charge density as the larger ones. Consequently, the capacitor with the smaller plate area stores less charge.

26.6 (*a*) The amount of charge remains the same; there are no other conductors to or from which charge carriers can go. (*b*) The charge carriers in the conducting slab always rearrange themselves so as to cancel the electric field inside the slab (see Section 24.5). To do so, they must set up an electric field of the same magnitude as that inside the capacitor, but in the opposite direction. The charge carriers in the slab rearrange themselves at the surfaces as shown in Figure 26.9*b*. For the electric field due to the charge carriers at the surfaces of the slab to be equal in magnitude to that caused by the capacitor plates, the magnitude of the surface charge density on the slab surfaces must be the same as that on the plates. Because the surface area of the slab (or at least the part of the slab that is in the electric field) is the same as that of the capacitor plates, the magnitude of the charge on each side of the slab is the same as that on the capacitor plates. (*c*) Zero, because the electric field inside the slab is zero. (*d*) It decreases. Between the capacitor plates and the surfaces of the slab, the electric field is the same as before, but inside the slab the electric field is now zero. The electrostatic work done on a unit charge as it is moved from one plate to the other is therefore less than it was before because it is zero inside the slab. Therefore the potential difference is decreased.

26.7 (*a*) No. Regardless of where the slab is inserted, the field inside it is always zero, while the field between the capacitor plates and the slab is equal to what it was before the slab was inserted. In the extreme case that the slab makes contact with one of the plates, the charge carriers on the plate move through the slab, all the way to its opposite side, as illustrated in Figure S26.7*a*, effectively reducing the plate separation. The magnitude of the electric field, however, remains the same. (*b*) See Figure 26.7*b*. Note that the slope of $V(x)$ on either side of the slab does not depend on the position of the slab.

26.8 The direction of the electric field created by a given source is defined to coincide with the direction of the electric force exerted on a *positive* charge carrier (see Section 23.2). Because electrons carry a negative charge, the force exerted on them is in a direction opposite the direction of the electric field.

Figure S26.7



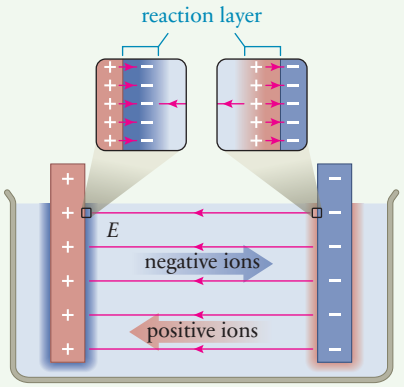
26.9 (*a*) The electric field is zero above the surface (so it has no direction) because the electric fields of the two surfaces cancel each other outside the dielectric, just like the fields of the two plates of a capacitor (see also Figure S23.14). (*b*) It points from the positively to the negatively charged surface—that is, opposite the direction of the electric field due to the capacitor plates.

26.10 (*a*) Zero, because the electric field due to the bound surface charge would be equal in magnitude and opposite in direction to that due to the free charge on the capacitor plates. In effect, the dielectric is like a conductor. (*b*) No. Suppose the bound charge accumulation at the surfaces becomes so great as to cancel the electric field due to the free charge on the capacitor plates, as in part *a*. Because the electric field inside the dielectric is zero, there is nothing that would cause the material to polarize even further.

26.11 The capacitor with the dielectric, because the battery must do *additional* work on the charge carriers to increase the magnitude of the charge on each plate above that of the capacitor without the dielectric.

26.12 The electric field is in the direction of the flow of positive ions. (The electric force exerted on a positive charge carrier is in the same direction as the electric field.) Wait! Shouldn't the field be in the opposite direction, given the charge on the terminals and the fact that the electric field lines point away from positive charge carriers? The key to reconciling these two facts is to realize that the mechanism of the voltaic cell occurs in the small layers of electrolyte near the electrodes. There the field is in the opposite direction and chemical reactions push charge carriers against the electric field. The result is the pattern of charge and electric fields shown (not to scale) in Figure S26.12.

Figure S26.12



26.13 The work is negative because the system under consideration comprises the molecules undergoing chemical reactions, not the electrons. The molecules do positive work on the electrons, which are part of the molecules' environment, so the environment does *negative* work on the molecules.

26.14 A. The charge on each capacitor follows from Eq. 26.1: $q = C\,V_{cap}$. Thus, for the same V_{cap}, the capacitor with the greater capacitance C holds the greater charge q.

26.15 (*a*) Using the result of Example 26.2, we have

$$A = \frac{Cd}{\epsilon_0} = \frac{(1.0\times10^{-6}\ \text{F})(50\times10^{-6}\ \text{m})}{8.85\times10^{-12}\ \text{C}^2/(\text{N}\cdot\text{m}^2)} = 5.6\ \text{m}^2.$$

This is the area of a small room! A 1.0-F capacitor would therefore have a surface area of 5.6×10^6 m^2—the size of a small town. (*b*) At the breakdown threshold, the potential difference across the capacitor is $V_{\text{cap}} = Ed = (3.0\times10^6\ \text{V/m})(50\ \mu\text{m}) = 150$ V. The charge on the capacitor is then $q = C\,\Delta V = (1.0\ \mu\text{F})(150\ \text{V}) = 1.5\times10^{-4}$ C. (*c*) This charge corresponds to $(1.5\times10^{-4}\ \text{C})/(1.6\times10^{-19}\ \text{C}) = 9.4\times10^{14}$ electrons.

26.16 The capacitance is given by the expression we obtained in Example 26.3. Substituting the values given in the checkpoint, we get

$$C = \frac{2\pi\epsilon_0\ell}{\ln(R_2/R_1)}$$

$$= \frac{2\pi[8.85\times10^{-12}\ \text{C}^2/(\text{N}\cdot\text{m}^2)](100\ \text{m})}{\ln[(2.0\times10^{-3}\ \text{m})/(2.0\times10^{-4}\ \text{m})]} = 2.4\ \text{nF}.$$

One nanofarad (1 nF) is 1×10^{-9} F, so this is a rather small capacitance. (A large capacitance is undesirable because it would allow charge carriers to "pile up" in the cable.)

26.17 (*a*) Example 26.4 gives an expression for the capacitance of a spherical capacitor. As R_2 approaches infinity, the R_1 in the denominator of this expression becomes negligible and so the capacitance reduces to

$$C = \lim_{R_2\to\infty}\left[4\pi\epsilon_0\frac{R_1R_2}{R_2-R_1}\right] = 4\pi\epsilon_0\frac{R_1R_2}{R_2} = 4\pi\epsilon_0R_1.$$

(*b*) Substituting values for ϵ_0 and the radius R_1, we get

$$C = 4\pi[8.85\times10^{-12}\ \text{C}^2/(\text{N}\cdot\text{m}^2)](2.5\ \text{m})$$

$$= 2.8\times10^{-10}\ \text{F}.$$

(*c*) The electric field is maximum at the surface of the dome. From Eq. 24.15 we know that the electric field at the surface of a charged sphere is

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{R^2} = \frac{1}{4\pi\epsilon_0}\frac{q}{R}\frac{1}{R} = V_R\frac{1}{R},$$

where V_R is the potential at the surface of the sphere and the potential is zero at infinity. Therefore, $V_R = ER$ and the charge stored on the dome is

$$q = CV_{\text{cap}} = CER$$

$$= (2.8\times10^{-10}\ \text{F})(3.0\times10^6\ \text{V/m})(2.5\ \text{m})$$

$$= 2.1\times10^{-3}\ \text{C}.$$

26.18 (*a*) The potential difference across the capacitor is

$$V_{\text{cap}} = Ed = (3.0\times10^6\ \text{V/m})(50\times10^{-6}\ \text{m}) = 150\ \text{V}.$$

The energy stored in the capacitor then follows from Eq. 26.4:

$$U^E = \tfrac{1}{2}(1.0\times10^{-6}\ \text{F})(150\ \text{V})^2 = 1.1\times10^{-2}\ \text{F}\cdot\text{V}^2.$$

Because 1 F = 1 C/V and 1 V = 1 J/C, we have $(1\ \text{F}\cdot\text{V}^2) = (1\ \text{J})$, so the energy stored in the capacitor is 11 mJ.

(*b*) The increase in gravitational potential energy is given by $\Delta U^G = mg\,\Delta y$, so $\Delta y = \Delta U^G/(mg) = (11\ \text{mJ})/[(2\ \text{kg})(10\ \text{m/s}^2)] = 0.6$ mm.

26.19 (*a*) The surface charge density on each plate is given by $\sigma = q/A$, so the magnitude of the field due to one plate is

$$E_{\text{plate}} = \frac{\sigma}{2\epsilon_0} = \frac{q}{2\epsilon_0A}.$$

Note that this is not the electric field between the capacitor plates. The field between the plates is twice as great because each plate contributes to that field. The force exerted by the electric field of one plate on the other plate is thus

$$F_{E\text{p}} = qE_{\text{plate}} = \frac{q^2}{2\epsilon_0A}.$$

(*b*) To increase the separation between the plates you must exert a force $\vec{F}$ of magnitude equal to the force exerted by the electric field on the plate calculated in part *a*. Because this force is constant, and because the force you exert must point in the same direction as the force displacement, we can write for the work you must do on the capacitor

$$W = F\Delta x_F = \frac{q^2\Delta x}{2\epsilon_0A}.$$

(*c*) The change in the electric potential energy is equal to the work done on the capacitor, which we determined in part *b*.

(*d*) The volume of the additional space is $A\,\Delta x$, so the energy stored in the electric field in this additional space is $u_E\,A\,\Delta x$. Because the electric field between the plates of the capacitor is twice the electric field of a single plate calculated in part *a*, the energy density of the electric field is

$$u_E = \tfrac{1}{2}\epsilon_0E^2 = \tfrac{1}{2}\epsilon_0\left(\frac{q}{\epsilon_0A}\right)^2 = \frac{q^2}{2\epsilon_0A^2},$$

so the additional energy is

$$\Delta U^E = u_E\,A\Delta x = \frac{q^2}{2\epsilon_0A^2}A\Delta x = \frac{q^2\Delta x}{2\epsilon_0A} = W.$$

26.20 (*a*) The energy stored in the capacitor is given by Eq. 26.4, so

$$U^E = \tfrac{1}{2}(10^{-4}\ \text{F})(300\ \text{V})^2 = 4.5\ \text{J}.$$

(*b*) The average power is given by the change in energy divided by the time interval over which the change takes place (Eq. 9.29):

$$P_{\text{av}} = \frac{4.5\ \text{J}}{1\times10^{-3}\ \text{s}} = 4.5\times10^3\ \text{W}.$$

This is a phenomenally large power (but it is delivered for only a very short amount of time)—much greater than the few watts that can be delivered by a typical battery. Because it requires a great amount of power to illuminate a large space for a short amount of time, all flash units use a capacitor that is charged by a battery between flash firings.

26.21 (*a*) Because $q_0 = q_{\text{free}} - q_{\text{bound}}$ (see Figure 26.29), we have $q_{\text{bound}} = q_{\text{free}} - q_0$. If we substitute Eq. 26.21, this becomes

$$q_{\text{bound}} = (\kappa - 1)q_0.$$

(*b*) The bound surface charge density on the dielectric is $\sigma_{\text{bound}} = q_{\text{bound}}/A$, where A is the area of the dielectric. Together with our answer to part *a*, this becomes

$$\sigma_{\text{bound}} = (\kappa-1)\frac{q_0}{A} = (\kappa-1)\epsilon_0E,$$

where $E = q_0/(\epsilon_0A)$ (Eq. 24.17).

26.22 (*a*) Because 1 V = 1 J/C and 1 J = 1 N·m, we have

$$\frac{\text{C}^2}{\text{N}\cdot\text{m}} = \frac{\text{C}^2}{\text{J}} = \frac{\text{C}}{\text{V}} = \text{F}.$$

(*b*) Similarly,

$$\text{F}\cdot\text{V}^2 = \frac{\text{C}}{\text{V}}\text{V}^2 = \text{C}\cdot\text{V} = \text{C}\left(\frac{\text{J}}{\text{C}}\right) = \text{J}.$$

26.23 Consider a cylindrical Gaussian surface of radius $r > R$. The electric flux through the surface of the cylinder is $\Phi = E(2\pi rL)$, where E is the electric field outside the insulation. The Gaussian surface encloses both bound and free charges. The enclosed bound charge is zero, however, because the surface encloses both the negative bound charge on the inside surface of the insulation and the positive bound charge on the outer surface. The free charge is still λL, so we obtain the same result as in Eq. 26.25 because $\kappa = 1$ for $r > R$.

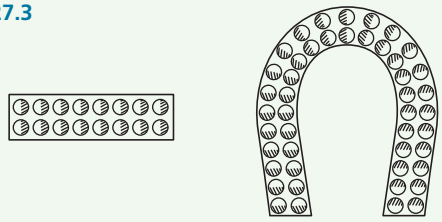
Chapter 27

27.1 No. Whether a pole is N or S is determined by the way the magnet orients itself with respect to Earth's North Pole. The experiment illustrated in Figure 27.3 then *shows* that like poles repel and unlike poles attract.

27.2 (*a*) Yes, because the charged object always pulls charge carriers of the opposite type toward itself, exerting an attractive force on the near side of the neutral object and a repulsive force on the far side. Because the electric force decreases with increasing distance, the magnitude of the attractive force is greater than that of the repulsive force and so the sum of the forces is attractive. (*b*) A north pole is induced in the clip on the left and a south pole on the right because the interaction between opposite magnetic poles is attractive.

27.3 (*a*) See Figure S27.3. (*b*) No poles because there are no faces where only one type of elementary magnet is exposed. (*c*) If you broke the ring in half and put the exposed faces near paper clips and other magnets, you would find that the faces of the cuts are magnetic poles.

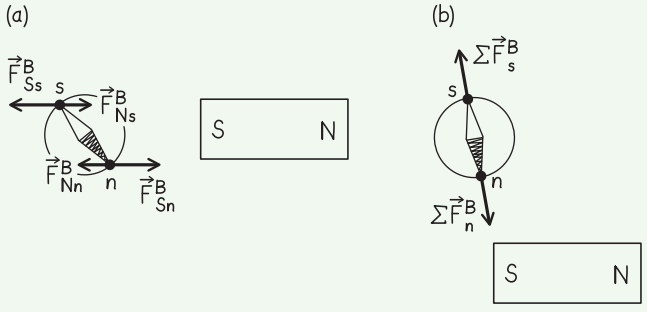
Figure S27.3



27.4 (*a*) By definition, the end that orients itself northward is the north pole of the needle. (*b*) The north pole of the magnet attracts the south pole of the needle and repels the north pole of the needle. Because the compass needle is free to rotate, the needle rotates so that its north-pointing end is as far away as possible from the north pole of the magnet. (*c*) No; if it were, it would repel the north pole of the compass needle. The earth's geographic North Pole is a magnetic south pole.

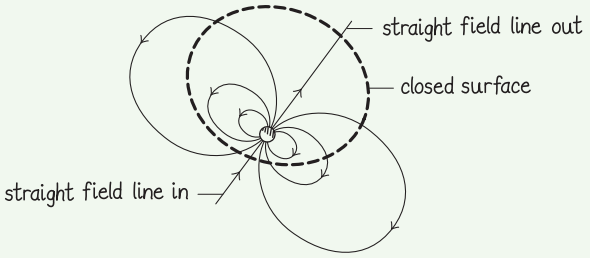
27.5 (*a*) The bar magnet's north pole attracts the needle's south pole and repels the needle's north pole. The two forces cause a torque on the needle that tends to align the needle with its north pole pointing away from the bar magnet's north pole. (*b*) Because the distance from the needle to the bar magnet's north pole is greater than the distance from the needle to the bar magnet's south pole, the magnitudes of the forces exerted by the north pole are smaller than those exerted by the south pole (Figure S27.5*a*). Consequently, the needle still aligns itself as in Figure 27.11*a*. (*c*) Once the needle has settled, we know that the torque on it must be zero. This means that the sum of the forces exerted by the poles of the bar magnet on each end of the needle must be aligned with the needle (Figure S27.5*b*).

Figure S27.5



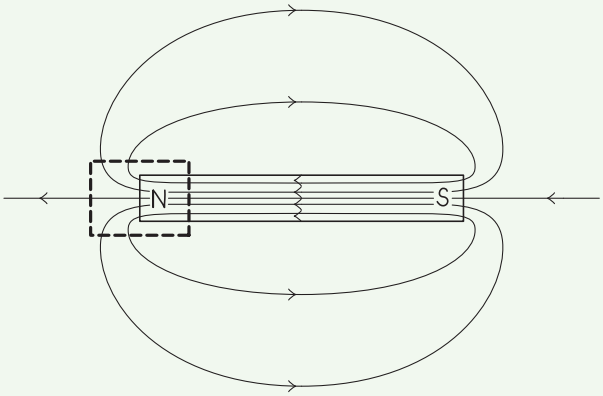
27.6 (*a*) Because elementary magnets have no physical extent, the magnetic field lines form loops that close on themselves (Figure S27.6). Because a particle cannot be cut in half, it must lie either inside or outside the closed surface. Consequently, each line that passes out through the closed surface must also pass back in through the surface. The straight field line entering at the bottom in Figure S27.6 is not an exception. It is balanced by the straight field line leaving at the top. The magnetic flux is zero. (*b*) No, because for this elementary magnet, too, each field line that leaves the surface must eventually reenter it. (*c*) Zero. Because the magnetic flux must be zero for each elementary magnet making up a bar magnet, the magnetic flux of the magnet must also be zero. (*d*) No. Because an elementary magnet cannot be divided, there is always a whole number of elementary magnets inside the closed surface, so the magnetic flux is still zero.

Figure S27.6



27.7 If we draw a closed surface around the north pole of the bar magnet (Figure S27.7), we see that all the field lines on the outside of the magnet are directed out of the closed surface. Because the magnetic flux through a closed surface must be zero, this means that the field lines inside the bar magnet point from the south pole to the north pole.

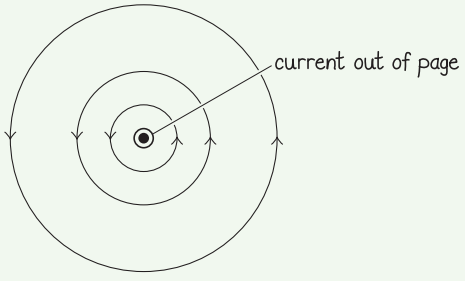
Figure S27.7



27.8 (*a*) No. If the rod were charged, the battery would have to carry an opposite charge, and charge carriers would start flowing back (see also Section 26.2). (*b*) No. Because it carries no surplus charge carriers, the rod cannot exert an electric force on another charged particle, and therefore there is no electric field due to the rod.

27.9 See Figure S27.9. If the needles align in a circular pattern, the magnetic field lines must be circular, curling in the direction that the north poles of the needles point.

Figure S27.9

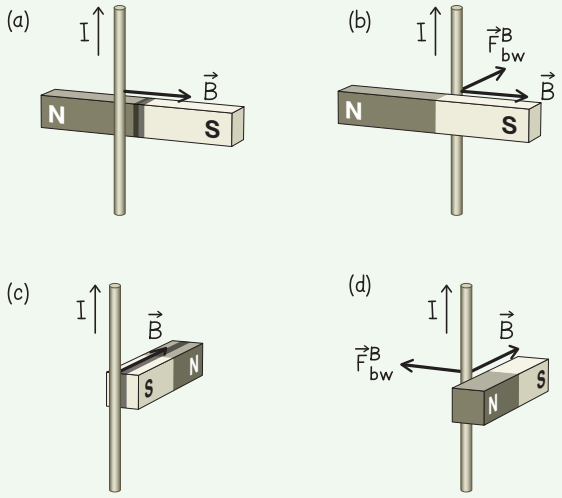


27.10 No. The circular magnetic field pattern generated by a current-carrying wire is unlike the field generated by any magnet or the elementary magnets of which any magnet is made. As we shall see in

the next chapter, however, it is possible to use a current to generate a magnetic field pattern that approximates that of a bar magnet.

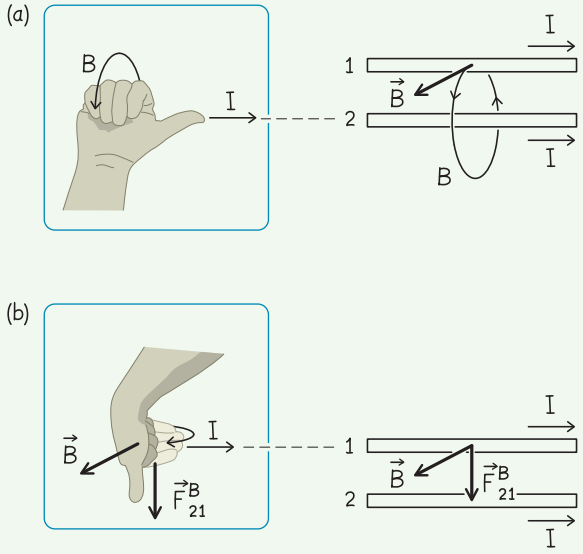
27.11 (*a*) Flipping the magnet in Figure 27.20*b* horizontally while keeping it behind the wire yields the situation in Figure S27.11*a*, with the magnetic field at the location of the wire pointing to the right. A field in the same direction can also be obtained by placing the magnet in front of the wire (Figure S27.11*b*). This is the same situation as shown in Figure 27.20*b*, but now seen from the back. Because the magnetic force exerted on the wire in Figure 27.20*b* is toward the front, it must be directed toward the back in Figure S27.11*b*. In other words, the force is directed *toward* the actual position of the magnet in Figure S27.11*a*. (*b*) Because the magnetic field lines loop from the north pole to the south pole outside the magnet (see Figure 27.13), the magnetic field at the wire points straight toward the south pole (Figure S27.11*c*). A magnetic field in the same direction can be obtained by placing the magnet to the right of the wire (Figure S27.11*d*). Comparing this situation with Figure 27.20*b*, you can conclude that the magnetic force exerted on the wire is now to the left. In other words, with the magnet in back of the wire as in Figure S27.11*c*, the magnet exerts a *sideways* force on the wire!

Figure S27.11



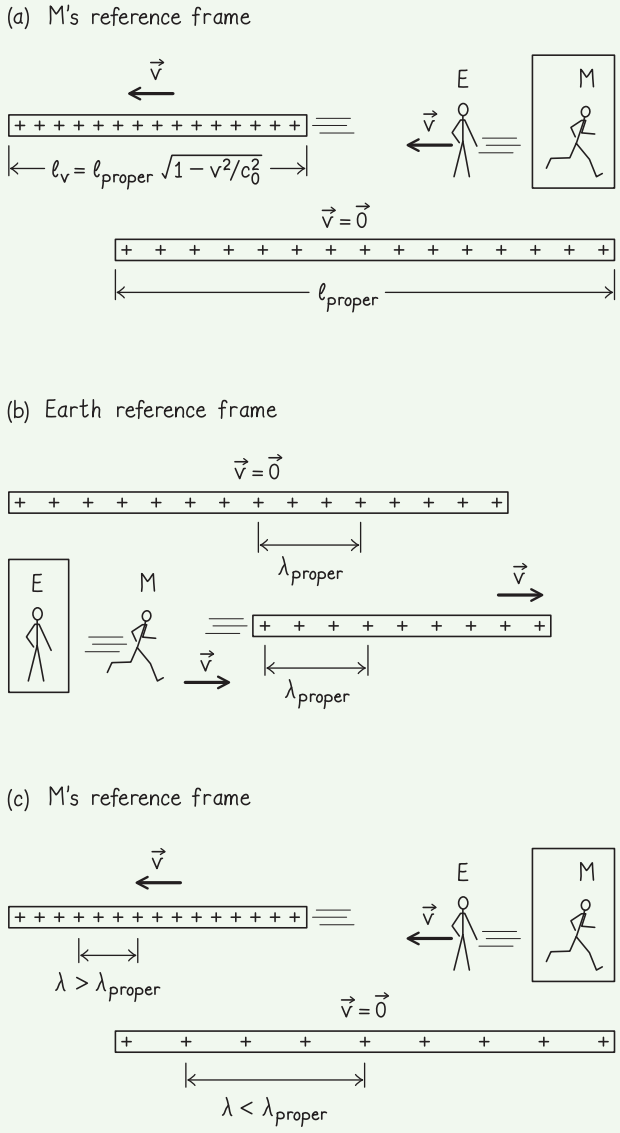
27.12 Reverse the direction of the current through rod 1 in Figure 27.23. As Figure S27.12*a* shows, at rod 1 the magnetic field due to the current through rod 2 is still out of the page. Placing the fingers of your right hand parallel to the current and curling them toward the magnetic field yield a downward-pointing thumb (Figure S27.12*b*), which means the magnetic force exerted by rod 2 on rod 1 is downward. The same reasoning tells you that the magnetic force exerted by rod 1 on rod 2 is upward. The two rods exert attractive forces on each other.

Figure S27.12



27.13 (*a*) From the point of view of observer M, the upper rod is moving to the left and the lower rod is at rest (Figure S27.13*a*). Consequently, the upper rod's length is contracted and the lower one has length ℓ_{proper}. Because both rods carry the same charge, the shorter rod—that is, the upper one—has the greater charge density. (*b*) In Figure S27.13*b*, the charge on the lower rod has been adjusted so that the charge density matches that of the upper rod. (Because the lower rod's length is contracted according to observer E, the charge it carries is now smaller than the charge carried by the upper rod.) According to observer M, the upper rod, which is moving, is shorter than the lower rod. Thus observer M sees the shorter upper rod carrying more charge than the longer lower rod and says $\lambda_{\text{upper rod}} > \lambda_{\text{proper}}$ and $\lambda_{\text{lower rod}} < \lambda_{\text{proper}}$ (Figure S27.13*c*).

Figure S27.13



27.14 (*a*) The positive ions moving to the left correspond to a current to the left. (*b*) The negative electrons moving to the right correspond to a current to the left. (*c*) Yes. The positive particle moving to the right corresponds to a current to the right, and so the situation in Figure 27.27*c* corresponds to two parallel rods carrying currents in opposite directions. As you saw in Example 27.1, the force between two such rods is repulsive, in agreement with the repulsive force exerted by the rod on the particle. (*d*) The arguments for the wire remain unchanged: It appears electrically neutral to observer E and positively charged to observer M. The positively charged wire exerts an attractive electric force on the negatively charged particle. (*e*) Yes. Because a negative particle moving to the right corresponds to a current to the left, the situation is identical to that of two rods carrying currents in the same direction. As you saw in Checkpoint 27.12, the force between such rods is indeed attractive.

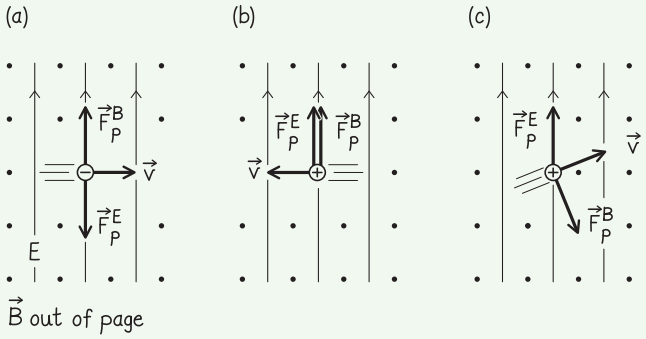
27.15 Example 27.2 shows that the current I is to the right in Figure 27.33. If this current is caused by negative charge carriers, the definition of current tells you that they must flow to the left.

27.16 (*a*) On the side through which the field enters, the field points into the cube but the area vector points outward, so $\theta = 180°$: $\Phi_B = AB\cos\theta = (1.0\ \text{m}^2)(1.0\ \text{T})\cos 180° = -1.0\ \text{Wb}$. The minus sign reflects the fact that the magnetic field points *into* the cube. (*b*) Because as many field lines enter the cube as leave it, the magnetic flux is zero.

27.17 Because the magnitudes of the charge on the proton and the electron are the same, Eq. 27.23 tells you that the ratio of the radii of the paths is equal to the ratio of the masses: $R_p/R_e = m_p/m_e = (1.67\times10^{-27}\ \text{kg})/(9.11\times10^{-31}\ \text{kg}) = 1.83\times10^3$.

27.18 (*a*) See Figure S27.18*a*. For negative charge carriers, $\vec{F}_p^E$ points in the direction opposite the direction of the electric field lines. Because q changes sign, $\vec{F}_p^B = q\vec{v}\times\vec{B}$ (Eq. 27.19) reverses direction as well. The magnitudes of the two forces are unchanged, and so they still add up to zero. (*b*) See Figure S27.18*b*. Reversing the direction of motion of the charge carriers does not affect $\vec{F}_p^E$, but it does reverse the direction of $\vec{v}$ and thus the direction of $\vec{F}_p^B = q\vec{v}\times\vec{B}$. Both forces therefore point in the same direction and so do not cancel. (*c*) See Figure S27.18*c*. Because $\vec{F}_p^B$ is always perpendicular to both the magnetic field and the velocity $\vec{v}$, the direction of $\vec{F}_p^B$ no longer lines up with that of $\vec{F}_p^E$, and so the two forces do not cancel.

Figure S27.18



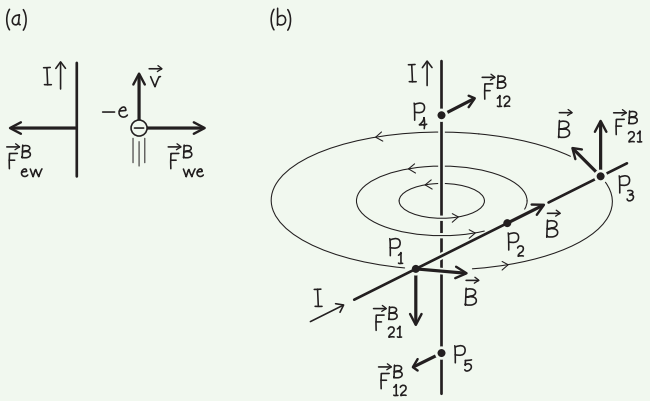
27.19 (*a*) The charge distribution is similar to that of a charged capacitor, so E is given by $E = |V_{RL}|/w$ (see Example 25.5), with the width w of the strip substituted for the separation between the positively and negatively charged sides. (*b*) Because the electric and magnetic forces exerted on the charge carriers are equal in magnitude, we have from Eq. 27.26 $v = E/B = |V_{RL}|/(wB)$. (*c*) From Eq. 27.16, $n = I/(A|q|v)$. Substituting the result from part *b*, we obtain $n = (IwB)/(A|qV_{RL}|)$. Because the charge carriers carry an elementary charge e and because $A = wh$, with h the height of the strip, we have $n = IB/(eh|V_{RL}|)$.

27.20 The wire has cylindrical symmetry and so the magnetic field spreads out only in two dimensions. The electric force seen by observer M must therefore decrease as $1/r$ (see Section 24.4). Because the electric force seen by observer M is the same as the magnetic force seen by observer E, that force must also decrease as $1/r$.

Chapter 28

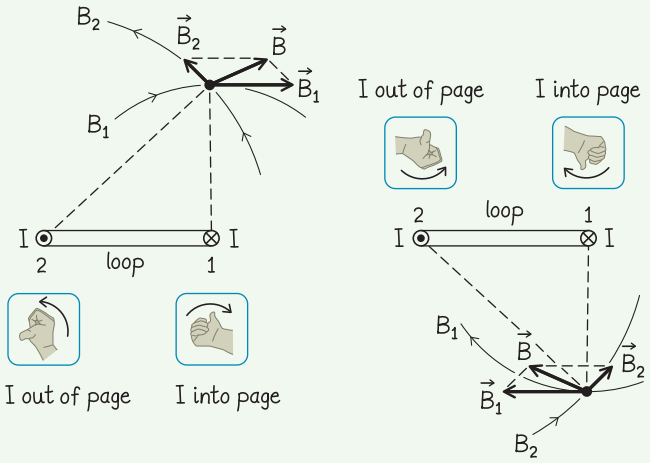
28.1 (*a*) See Figure S28.1*a*. (*b*) The two particles exert no magnetic forces on each other because it takes a moving charged particle to detect the magnetic field of another moving charged particle. (*c*) See Figure S28.1*b*. The forces exerted on the horizontal wire cause a torque that tends to align that wire with the vertical one. Conversely, the forces exerted on the vertical wire cause a torque in the opposite direction, tending to align that wire with the horizontal one.

Figure S28.1



28.2 See Figure S28.2.

Figure S28.2

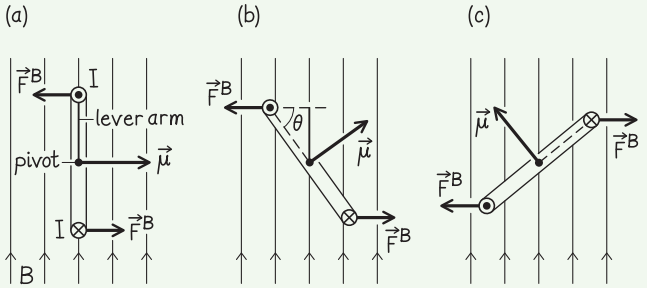


28.3 Comparing Figures 28.6*b* and 28.7, you see that the positively charged ring has a magnetic field similar to the field of a bar magnet, with its north pole up. The negatively charged ring thus has its north pole down, and this north pole is directly above the north pole of the positive ring. Because two north poles repel each other, the interaction is repulsive.

28.4 No. By definition the electric field points from positive charge carriers to negative charge carriers (see Section 23.2). The electric field along the axis passing through the poles of an electric dipole therefore points from the positive end to the negative end. The electric dipole moment, however, by definition points in the opposite direction (see Section 23.6).

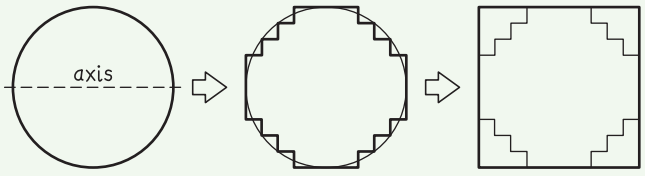
28.5 (*a*) Because the magnetic field remains perpendicular to the horizontal sides of the loop, you know from Eq. 27.4 that the magnitude of the magnetic force exerted on each side is $F^B_{\max} = |I|\ell B$, where ℓ is the length of each side of the loop. Because none of these quantities changes, the magnitude of the force exerted on the horizontal sides stays the same. (*b*) See Figure S28.5*a* and *b*. The lever arm of the force exerted on the horizontal sides become smaller as the loop rotates, and so the torque caused by these forces decreases. (*c*) In the initial position, the vertical sides are parallel to the magnetic field, which means the magnetic force exerted on them is zero (Eq. 27.7, $F^B = |I|\ell B \sin\theta$ with $\theta = 0$). After the loop has rotated 90°, the vertical sides are perpendicular to $\vec{B}$ and thus subject to an outward magnetic force. For $0 < \theta < 90°$, the magnitude of the force is again given by Eq. 27.7. As you can verify using the right-hand force rule, the forces exerted on the vertical sides are directed outward along the rotation axis and so cause no torque. (*d*) See Figure S28.5*c*. Because the top and bottom of the loop are now reversed, the direction of the torque reverses.

Figure S28.5



28.6 See Figure S28.6, where the circular loop is approximated by a series of vertical and horizontal segments. The vertical segments experience no force, but the horizontal ones do. All the horizontal segments add up to the same length as two sides of a square straddling the circle, and thus Eq. 27.4, $F^B_{\max} = |I|\ell B$, tells you that the magnitude of the magnetic force exerted on the horizontal segments is the same as the force exerted on the horizontal side of the square loop. Some horizontal segments are closer to the rotation axis, however, and so the lever arms of the forces acting on these closer segments are smaller than the lever arms of the forces acting on the square loop. Therefore the torque on the circular loop is smaller than that on the square loop.

Figure S28.6



28.7 With the magnetic field on the left stronger than that on the right, $F^B_{\text{left}} > F^B_{\text{right}}$. Therefore the vector sum of the forces exerted on these two sides is nonzero and points upward. The effect is a clockwise rotation due to the torque and an upward acceleration due to the upward vector sum of the forces.

28.8 Increasing the current increases the magnitude of the magnetic field, and so the line integral increases.

28.9 (*a*) If the direction of the current is reversed, the direction of the magnetic field is reversed, and so the algebraic sign of the line integral is reversed. (*b*) Because the value of the line integral of the magnetic field around a closed path does not depend on the position of the wire inside the path, the second wire by itself gives rise to the same line integral as the first wire. Adding the second wire thus doubles the value of the line integral. (*c*) Reversing the current though the second wire flips the sign of the line integral. If its value was C, it is $-C$ after the current is reversed. The sum of the two line integrals is thus $C + (-C) = 0$.

28.10 No. If the path tilts, you can break it down into arcs, radial segments, and *axial* segments, which are segments parallel to the wire. Because the magnetic field is perpendicular to the axial segments, they don't contribute to the line integral. For the same reason, of course, the radial segments contribute nothing. So the line integral again is that along all the arcs, which is equivalent to a single complete circular path.

28.11 (*a*) Reversing the direction of the current through wire 1 changes the sign of its contribution from negative to positive, so adding the contributions of wires 1 and 3 yields a positive value. (*b*) The path does not encircle wire 2, so reversing the current through wire 2 does not affect the answer to Exercise 28.2. (*c*) Reversing the direction of the path inverts the signs of all the contributions to the line integral, but because the integral is zero, the answer doesn't change.

28.12 See Figure S28.12. Inside the wire, use Ampèrian path 1 of radius $r < R$. For this path,

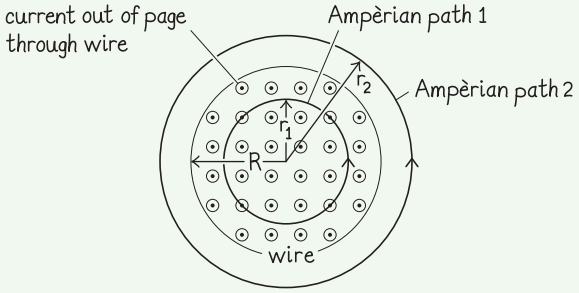
$$\oint \vec{B}\cdot d\vec{\ell} = B_{\text{inside}} \oint d\vec{\ell} = B_{\text{inside}}\, 2\pi r,$$

but the path encircles only part of the cross section and so encircles only part of the current I through the wire. The area of the cross section of the wire is πR^2, and the cross-sectional area of Ampèrian path 1 is πr^2, making the fraction of the wire cross section enclosed by the path $(\pi r^2)/(\pi R^2) = r^2/R^2$. Thus, the right side of Ampère's law is $\mu_0 I_{\text{enc}} = \mu_0(r^2/R^2)I$. Ampère's law thus yields $B_{\text{inside}}(2\pi r) = \mu_0(r^2/R^2)I$, or $B_{\text{inside}} = \mu_0 Ir/2\pi R^2$. Outside the wire, use Ampèrian path 2 in Figure S28.12. For this path,

$$\oint \vec{B}\cdot d\vec{\ell} = B_{\text{outside}} 2\pi r.$$

Because the path encircles all of the current I through the wire, $\mu_0 I_{\text{enc}} = \mu_0 I$, and so the magnetic field outside the wire is the same as that for a long thin wire: $B_{\text{outside}} = \mu_0 I/2\pi r$.

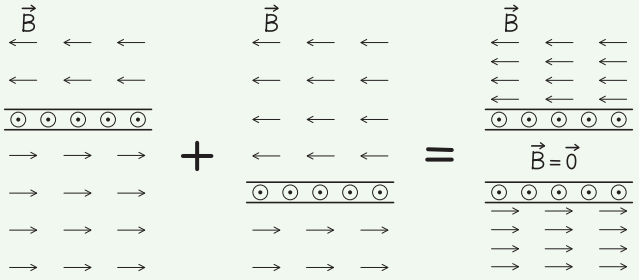
Figure S28.12 Cross section through a current-carrying wire



28.13 (*a*) Figure S28.13 on the next page shows a head-on view of the magnetic fields of the two sheets separately and their sum. The magnetic fields add up to zero between the sheets. Below and above the sheets, the magnetic field points to the left and its magnitude is twice that of a single sheet: $2(\frac{1}{2}\mu_0 K) = \mu_0 K$. (*b*) With the current through the lower sheet reversed, its magnetic field is also reversed.

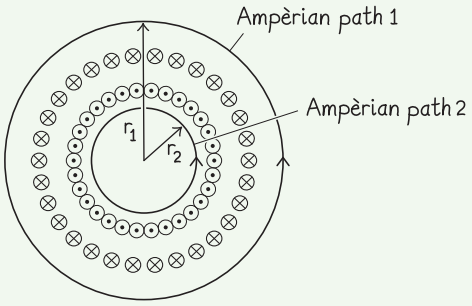
Now the magnetic fields add up to zero below and above the sheets. Between them the magnetic field is to the right and its magnitude is twice that of a single sheet: $\mu_0 K$.

Figure S28.13



28.14 Outside the toroid, you use Ampèrian path 1 in Figure S28.14. The left side of Ampère's law is still given by Eq. 28.7, but the path now encircles both the current upward through the inside of each winding and the current downward through the outside of each winding. The encircled current is thus $NI - NI = 0$, making B outside the toroid zero. Inside the toroid's inner radius, use Ampèrian path 2. This path encloses no current, so here, too, $B = 0$.

Figure S28.14



28.15 For this wire the integration in Example 28.6 extends from $x = 0$ to $x = +\infty$:

$$B = \frac{\mu_0 I d}{4\pi}\int_0^{+\infty}\frac{dx}{[x^2+d^2]^{3/2}}$$

$$= \frac{\mu_0 I d}{4\pi}\left[\frac{1}{d^2}\frac{x}{[x^2+d^2]^{1/2}}\right]_{x=0}^{x=+\infty} = \frac{\mu_0 I}{4\pi d}.$$

28.16 (*a*) Using our result from Example 28.7 with $\phi = 2\pi$:

$$B = \frac{\mu_0 I(2\pi)}{4\pi R} = \frac{\mu_0 I}{2R}.$$

(*b*) The outer arc spans an angle $\phi = \pi/2$ and has a radius $2R$. Using our result from Example 28.7 again, it contributes a magnetic field

$$B_{\text{outer}} = \frac{\mu_0 I(\pi/2)}{4\pi(2R)} = \frac{\mu_0 I}{16R}.$$

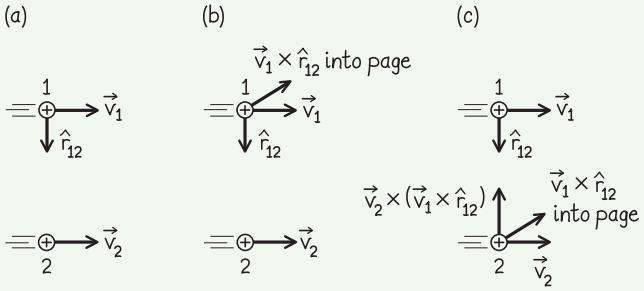
The inner arc also spans an angle of $\phi = \pi/2$, but it has radius R, and the current through it runs in the opposite direction so the magnetic field due to this arc is directed into the page. For this arc, the magnitude of the magnetic field is

$$B_{\text{inner}} = \frac{\mu_0 I(\pi/2)}{4\pi(R)} = \frac{\mu_0 I}{8R}.$$

The two straight segments do not contribute because they point straight toward P. Because B_{inner}, which points into the page, is greater than B_{outer}, which points out of the page, the magnetic field $\vec{B} = \vec{B}_{\text{inner}} + \vec{B}_{\text{outer}}$ is into the page rather than out of the page. The magnitude of the magnetic field is $B = \mu_0 I/8R - \mu_0 I/16R = \mu_0 I/16R$.

28.17 (*a*) The direction of the magnetic force is given by the vector products in Eq. 28.23. Figure S28.17*a* shows the vectors that appear in this expression. You must first evaluate $\vec{v}_1 \times \hat{r}_{12}$. The direction of the resulting vector is obtained by curling the fingers of your right hand from $\vec{v}_1$ to $\hat{r}_{12}$, which tells you $\vec{v}_1 \times \hat{r}_{12}$ points into the plane of the page (Figure S28.17*b*).

Figure S28.17



Next evaluate the vector product between $\vec{v}_2$ and the vector product you just found. The right-hand vector product rule tells you this yields a vector that points upward from proton 2 to proton 1. Thus, $\vec{F}_{12}^{B}$ points upward. Applying the same procedure to obtain $\vec{F}_{21}^{B}$, you find that it points downward, telling you that the magnetic interaction is attractive, like that of two parallel current-carrying wires.

(*b*) Because the angles involved in the vector products in Eq. 28.23 are all 90°,

$$F_{12}^{B} = \frac{\mu_0}{4\pi}\frac{q_1 q_2}{r_{12}^2} v_1 v_2.$$

The magnitude of the electric force is given by Coulomb's law:

$$F_{12}^{E} = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r_{12}^2}.$$

The ratio is

$$\frac{F_{12}^{B}}{F_{12}^{E}} = \frac{\left(\frac{\mu_0}{4\pi}\right)\frac{q_1 q_2}{r_{12}^2} v_1 v_2}{\left(\frac{1}{4\pi\epsilon_0}\right)\frac{q_1 q_2}{r_{12}^2}} = \mu_0 \epsilon_0 v_1 v_2.$$

Substituting numerical values, you find that this ratio is 10^{-6}.

Chapter 29

29.1 The magnetic force exerted on the charge carriers points in the opposite direction—that is, upward rather than downward. As a result, positive charge accumulates at the upper end of the rod rather than the lower end.

29.2 Yes, in positions (*b*) and (*d*) in Figure 29.4. The only difference between moving the field and moving the loop is the observer's frame of reference. It would not make any sense for an observer in one reference frame to observe a current and an observer in another reference frame to see no current.

29.3 No, because magnetic forces are exerted only on moving charged particles, not on stationary ones.

29.4 There are two electric fields in Figure 29.10: the "regular" electrostatic field due to the charge separation and the electric field that accompanies the changing magnetic field due to the moving magnet. The charge distribution in the rod is in mechanical equilibrium when these two fields have the same magnitude and opposite directions, producing zero electric field inside the stationary rod, as we would expect.

29.5 Situation 1: The induced current must be the same as if the loop were moving to the left with the magnet stationary (the same relative motion). The upward component of the magnetic field exerts a force on a positive charge carrier that is directed into the page (toward the back of the loop). Because at this point in the magnet's motion the magnetic field is stronger on the left side of the loop than on the right side, the induced current is clockwise.

Situation 3: The induced current must be the same as if the loop were moving downward with the magnet stationary. The upward component of the magnetic field does not contribute to the magnetic force (because it's parallel to the motion of the charge carriers). The component pointing radially outward (looking down on the loop) produces forces that drive a clockwise current around the loop (viewed from above).

29.6 (*a*) In Example 29.4 we saw that the induced current is clockwise as viewed from above in Figure 29.12*b*. Inside the loop $\vec{B}_{\text{ind}}$ points down, and outside the loop $\vec{B}_{\text{ind}}$ points up. Remember that $\vec{B}_{\text{ind}}$ is stronger inside the loop than outside and that the direction of the force exerted by $\vec{B}_{\text{ind}}$ on the south pole is opposite the direction of $\vec{B}_{\text{ind}}$. Thus, the force exerted on the magnet by $\vec{B}_{\text{ind}}$ pushes the magnet upward, opposing the downward motion of the magnet that induced the current and field $\vec{B}_{\text{ind}}$.
(*b*) If Lenz's law indicated that the induced current and field should add to the change that produced them, the force exerted on the magnet by $\vec{B}_{\text{ind}}$ would accelerate the magnet downward more rapidly (and in the process increase the induced current).

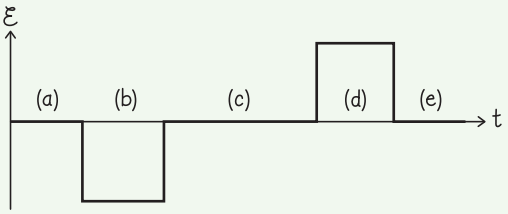
29.7 (*a*) Zero. Once the loop is completely in the magnetic field, the magnetic flux enclosed by the loop is no longer changing. The loop's motion ceases to induce current, and the force that resists the motion vanishes, which means no further work is required to keep the loop moving at constant speed. (*b*) Positive. Once the right edge leaves the field, the magnetic flux enclosed by the loop begins to decrease. Lenz's law says that a force resists the magnetic flux decrease and therefore resists the loop's motion. Thus, the loop must be pulled out. The pulling force you exert points in the same direction as the motion, so the work done is positive.

29.8 More work must be done to move a magnet toward a closed conducting loop because the motion induces a current in the loop, and the work done on the magnet must provide the energy for this current. Because there cannot be a current through a rod, moving a magnet toward a rod induces only a static charge separation, which requires less work.

29.9 (*a*) The magnitude decreases through a small region around P because the magnet is moving away from P. It remains unchanged through a small region around Q because the magnet is directly overhead. It increases through a small region around R because the magnet is moving toward R. (*b*) The direction of any eddy current is such that the magnetic field associated with the current opposes the change in magnetic flux. A counterclockwise current (viewed from overhead) creates a field that points upward. Therefore the eddy current around P is counterclockwise and that around R is clockwise. The magnetic flux through a small region around Q remains unchanged, so there are no eddy currents around Q.

29.10 See Figure S29.10 (with magnetic flux out of the page in Figure 29.4 chosen to be positive). The induced emf is zero when the loop is completely outside (*a* and *e*) or completely inside (*c*) the field. The emf has the same constant value when it is moving into (*b*) or out of (*d*) the field. (The sign of the emf is different in *b* and *d*, but the magnitude is the same because $|\Delta\Phi_B/\Delta t|$ depends only on the speed of the loop and the magnitude of $\vec{B}$, both of which are constant.)

Figure S29.10



29.11 Although the solenoid orientation looks the same, the two sides have interchanged their positions. Because the two sides have changed positions, the sense of circulation has changed too. The emf thus has changed sign too.

29.12 (*a*) There is no change in the magnetic flux and thus no electric field. (*b*) They look like counterclockwise-pointing circular loops. By Lenz's law, the magnetic field of the induced current must oppose the decrease in the magnitude of the magnetic field. Therefore the induced current is counterclockwise and the electric field lines must form circles and the electric field points in the direction opposite the direction it has when the magnetic field is increasing.

29.13 For $r > R$, the left side of Eq. 29.17 is still given by Eq. 1 in Example 29.7, but the right side of Eq. 29.17 is $\pi R^2\, dB/dt$. This gives

$$2\pi r\, E = \pi R^2 \frac{dB}{dt}$$

$$E = \frac{R^2}{2r}\frac{dB}{dt}.$$

Substituting $dB/dt = 0.050$ T/s, $R = 0.20$ m, and $r = 0.30$ m into this expression yields

$$E = \frac{(0.20\text{ m})^2(0.050\text{ T/s})}{2(0.30\text{ m})} = 3.3\times10^{-3}\text{ N/C}.$$

29.14 You can calculate the induced emf from the inductance of the solenoid (Example 29.8) and Eq. 29.19:

$$L = \frac{(4\pi\times10^{-7}\text{ T}\cdot\text{m/A})(2760)^2\pi(50\times10^{-3}\text{ m})^2}{0.60\text{ m}} = 0.13\text{ H}$$

$$|\mathscr{E}_{\text{ind}}| = L\frac{dI}{dt} = (0.13\text{ H})(1.0\times10^{-1}\text{ A/s}) = 13\times10^{-3}\text{ V}.$$

29.15

$$u_B = \frac{B^2}{2\mu_0} = \frac{(1.0\text{ T})^2}{2(4\pi\times10^{-7}\text{ T}\cdot\text{m/A})} = 4.0\times10^5\text{ J/m}^3$$

$$u_E = \tfrac{1}{2}\epsilon_0 E^2 = \tfrac{1}{2}(8.85\times10^{-12}\text{ C}^2/\text{N}\cdot\text{m}^2)(1.0\text{ V/m})^2 = 4.4\times10^{-12}\text{ J/m}^3.$$

The ratio of u_B to u_E is 9.0×10^{16}.

Chapter 30

30.1 (*a*) Yes. The wire is intercepted by the surface three times, but the directions of the intercepts are different. Because the magnetic field points in the same direction as the integration direction along the closed path, we can use the right-hand current rule (see the Procedure box in Chapter 28, point 4) to determine that intercepts 1 and 3 are positive and intercept 2 is negative. The direction of the current at intercepts 1 and 3 is from inside to outside; at intercept 2 it is from outside to inside. We can expect current intercepted going from inside to outside to contribute oppositely to current intercepted going from outside to inside. From Figure 30.1 we see that current intercepted going from inside to outside through surface B is positive, so the current intercepted by the surface is $+I$ at 1 and 3 and $-I$ at 2, for a total of I, which is equal to the current encircled by the path. (*b*) Yes. The current intercepted by the surface is I: $+I$ for the top intercept, $-I$ for the middle intercept, and $+I$ for the bottom intercept. Again, the current intercepted by the surface is equal to the current encircled by the path.

30.2 (*a*) While the capacitor is charging, there is a current through the wire leading to or from the capacitor, and the electric field between the plates is nonzero and changing. (*b*) Once the capacitor is fully charged, the current through the wire drops to zero. There is still a nonzero electric field between the plates, but that field is no longer changing.

30.3 Although the electric field direction does not change, the direction of the *change* $\Delta\vec{E}$ in the electric field is reversed because the electric field is now decreasing. Taking the direction of $\Delta\vec{E}$ as the "current," you see that the direction of the magnetic field is reversed from what it was when the unit was charging, and the right-hand current rule applies.

30.4 The magnetic dipole moment of the neutron indicates that it must have an internal structure consisting of charged particles that form current loops. Indeed, in Chapter 7 we learned that the neutron consists of one up and two down quarks (Figure 7.17). These quarks do indeed carry charge.

30.5 You can estimate v by comparing the distance the particle moves in the time interval between parts *b* and *c* of Figure 30.8 with the distance the kinks have moved in that interval as measured from the point where they originate, which is at the center of each panel. The kinks move roughly twice as far as the particle, which means $v \approx 0.5c$.

30.6 The electric field is changing at the kinks and in the region between the kinks and the particle. In these regions there is a magnetic field. Beyond the kinks, the magnetic field is zero because the electric field is static.

30.7 (*a*) In a time interval equal to the period T, the wave travels a distance equal to its wavelength. Because the wave travels horizontally, you should measure the distance traveled along the horizontal axis through the electric field pattern. Between $t = 0$ and $t = T/2$ the wave advances half a wavelength. The distance that the rightmost field line advances in this time interval is 16 mm, so the wavelength is 32 mm. (*b*) From Eq. 16.10, $f = c/\lambda = 9.4 \times 10^9$ Hz. (*c*) From Eq. 15.2, $T = 1/f = 1.1 \times 10^{-10}$ s, or 0.11 ns.

30.8 (*a*) Figure 30.15*a* shows the bottom electric field pattern of Figure 30.14, so the motion of the particles must be the same in those two cases. From Figure 30.14 you see that at $t = \frac{3}{4}T$, the positively charged particle is moving up and the negatively charged particle is moving down, which means the current direction is up. (Although the dipole moment is zero at this instant because the particles are in the same location, the current is not zero because they are both moving. The current is instantaneously zero when the particles are at the two extremes of their travel paths and thus have zero velocity for an instant.) (*b*) Yes, because the magnetic field lines closest to the dipole are in the proper direction for the magnetic field of an upward current: Immediately to the right of the dipole, the magnetic field lines point into the page; immediately to the left of the dipole, they point out of the page.

30.9 Parallel, because parallel to the polarization means parallel to the electric field. From Section 23.4, you know that an electric field exerts a force on charged particles (see Figure 23.15, for instance). Only electric forces with a component along the axis of the antenna can accelerate the charge carriers in the antenna in a direction that produces current (along the antenna's length), so an electric field with a component parallel to the antenna is needed.

30.10 Whether between the plates or outside the capacitor, the magnetic field accompanies a change in the electric field that points left (the electric field is decreasing). From Section 30.1, you know that the "current" you must use with the right-hand current rule for determining the direction of $\vec{B}$ also points left. If you point the thumb of your right hand to the left along the axis in Figure 30.24, your curled fingers point into the page above the axis of the capacitor and out of the page below the axis. This means $\vec{B}$ is into the page at both P and S.

30.11 (*a*) Use a surface that lies between the plates, like the one in Figure 30.28 but with a radius $r < R$. Then follow the approach used in Example 30.5, substituting r for R in Eq. 1:

$$B = \frac{\mu_0\epsilon_0 r}{2}\frac{dE}{dt}.$$

To determine E, use Eq. 1 from Example 26.2 with $A = \pi R^2$ (because A is the area of each plate, not the area of the surface bounded by the path of integration). This substitution gives you

$$B = \frac{\mu_0\epsilon_0 r}{2}\frac{d}{dt}\left(\frac{q}{\epsilon_0\pi R^2}\right) = \frac{\mu_0 r}{2\pi R^2}\frac{dq}{dt}.$$

You know that dq/dt is the current, so you can write $B = \mu_0 Ir/(2\pi R^2)$. Thus, the magnetic field between the plates is smaller closer to the axis of the plates. (*b*) To the right of the right plate (and, of course, to the left of the left plate), the magnetic field is simply that of a current-carrying wire at a distance r from the wire. The result we obtained in Example 28.3 gives you $B = \mu_0 I/(2\pi r)$.

30.12 Equation 30.10 relates the electric field to its source charge distributions and thus needs no modification. The right side of Eq. 30.11 is zero because there are no magnetic monopoles. If monopoles having a $1/r^2$ relationship did exist, Eq. 30.11 would have to be analogous to Eq. 30.10. This means the right side must be proportional to m_{enc} rather than zero. You would expect electric fields to form loops around monopole currents, in the same way that magnetic fields form loops around currents, so the right side of Eq. 30.12 must gain a term proportional to dm/dt. Equation 30.13 needs no modification because $\oint \vec{B}\cdot d\vec{\ell}$ for a static distribution of monopoles should be zero, just as $\oint \vec{E}\cdot d\vec{\ell}$ for a static charge distribution is zero.

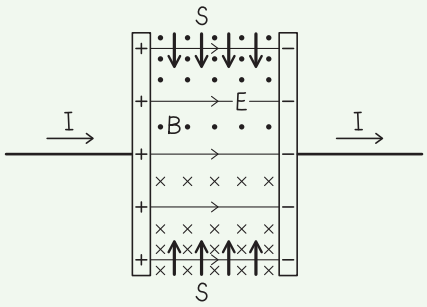
30.13 In vacuum, the line integral of the electric field is proportional to the time rate of change of magnetic flux through the surface bounded by the path (Eq. 3), and the line integral of the magnetic field is proportional to the time rate of change of electric flux through the surface bounded by the path (Eq. 4). To satisfy these relationships, as shown in Figure 30.30*e* and *f*, the magnetic field and electric field must be perpendicular to each other. The magnetic field must form loops around the electric field, and the electric field must form loops around the magnetic field.

30.14 When an electromagnetic wave enters a dielectric medium, the electric field of the wave accelerates charged particles in the medium back and forth with the frequency of the wave. The accelerated particles radiate electromagnetic waves that propagate with the same frequency as the incoming wave. The electromagnetic wave in the medium is the combination of all of these waves and hence has the same frequency, $f = c_0/\lambda = (3.00 \times 10^8 \text{ m/s})/(600 \text{ nm}) = 5.00 \times 10^{14}$ Hz. The wavelength in the dielectric medium decreases because the speed of the wave in the medium is smaller than in vacuum:

$$\lambda_{\text{new}} = \frac{c}{f} = \frac{c_0}{f\sqrt{\kappa}} = \frac{\lambda}{\sqrt{\kappa}} = \frac{600 \text{ nm}}{\sqrt{1.30}} = 526 \text{ nm}.$$

30.15 The electric field points from one plate to the other, and the magnetic field forms loops around the electric field. (Figure 30.4 represents this situation.) The vector product of these two vectors points radially inward, toward the capacitor axis (Figure S30.15). The Poynting vector thus represents the flow of energy into the region between the capacitor plates, where energy is being stored in the electric field. This makes sense because the energy density associated with the electric field inside the cylindrical surface defined by the capacitor plates is increasing. So there must be a flow of electromagnetic energy into this region.

Figure S30.15



30.16 The SI units of μ_0 are T · m/A; those of the electric field N/C and of the magnetic field T. Using Eq. 30.36, I thus get for the SI units of the Poynting vector:

$$\frac{[\text{N/C}][\text{T}]}{\text{T}\cdot\text{m/A}} = \frac{\text{N}\cdot\text{A}}{\text{C}\cdot\text{m}} = \frac{\text{N}\cdot\text{A}}{\text{A}\cdot\text{s}\cdot\text{m}} = \frac{\text{N}}{\text{m}\cdot\text{s}} = \frac{\text{J}}{\text{m}^2\cdot\text{s}} = \frac{\text{W}}{\text{m}^2}.$$

Chapter 31

31.1 (*a*) Yes. In practice, the light and thermal energy travel away from the bulb, so you need either a flexible definition of the system or a container capable of keeping the light and thermal energy in a well-defined volume. (*b*) Yes. (*c*) The energy to produce light and thermal energy comes from the electric potential energy associated with the potential difference between the battery terminals. This energy comes from chemical reactions taking place inside the battery.

31.2 (*a*) Yes. The charged capacitor is the source, and the bulb is the load. (*b*) The bulb glows during the brief time interval when the capacitor is discharging. Electrons flow from the negatively charged capacitor plate through the connecting wires and the bulb to the positively charged plate until the capacitor plates are no longer charged. (*c*) Electric potential energy stored in the charged capacitor is converted to light and thermal energy in the bulb.

31.3 These answers are appropriate for a flashlight bulb connected to four D batteries that are fresh. For different batteries and different bulbs, the times may vary somewhat. (*a*) The current should stay constant for many minutes. (*b*) After 24 hours, the batteries are depleted and the current is zero.

31.4 (*a*) Gains electric potential energy. (*b*) Loses the same amount of electric potential energy. (*c*) In the load (they lose very little energy when flowing through the wires).

31.5 (*a*) That resistance is proportional to potential difference tells you that A has the greater resistance. (*b*) The greater resistance of bulb A means that a greater potential difference is required to obtain a given current in A. Thus, the same potential difference across both bulbs produces a smaller current in A than in B.

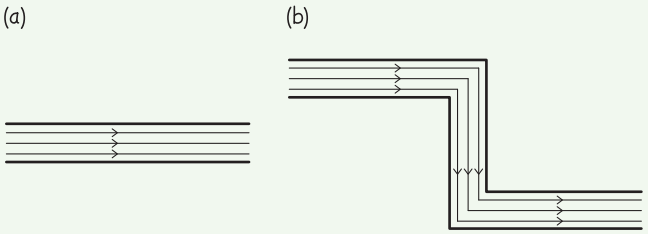
31.6 Smaller than. When the circuit load consists of the parallel combination of two bulbs, the current out of the battery is greater than the current out of the battery when the load consists of a single bulb. (To be exact, the current is twice as great with the parallel combination of bulbs as with the single bulb, as shown in Example 31.3.) The potential difference across the load is the same in both cases. Therefore, the resistance of the parallel combination of two bulbs is half the resistance of the single bulb.

31.7 (*a*) One contact of A is connected to the positive terminal of battery 1, and the other contact is connected to the negative terminal of battery 2. The negative terminal of battery 1 is connected to the positive terminal of battery 2. Therefore the potential difference across A equals the potential difference across the two-battery combination, which is the sum of the potential differences across each battery, or 18 V. (*b*) One contact of bulb B is connected to the positive terminal of battery 1, and the other contact of bulb B is connected to the negative terminal of the same battery. Therefore, the magnitude of the potential difference across B must be 9 V. (*c*) Bulb A glows more brightly because there is a greater potential difference across it.

31.8 Because the electric field is uniform throughout the rod, the magnitude of the electric field is equal to the magnitude of the potential difference across the rod divided by the length of the rod, $E = |V_{12}|/\ell$ (see Eq. 1 in Example 25.5).

31.9 See Figure S31.9. The electric field lines are parallel to the sides of the conductors everywhere. The electric field line density is also uniform, representing the uniform magnitude of the electric field.

Figure S31.9

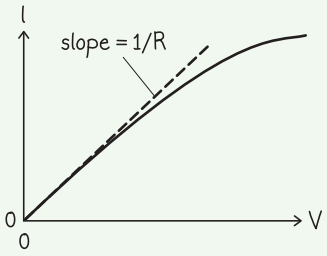


31.10 (*a*) You know from the text that $V_{\text{wide}} = V_{\text{narrow}}/4$. You also know that the magnitude of the potential difference across the entire conductor is 9 V. Therefore $V_{\text{wide}} + V_{\text{narrow}} = V_{\text{wide}} + 4V_{\text{wide}} = 9$ V, so $V_{\text{wide}} = 9\text{ V}/5 = 1.8$ V. (*b*) $V_{\text{narrow}} = 9\text{ V} - 1.8\text{ V} = 7.2$ V.

31.11 (*a*) Yes. The electric field does work on the electrons to accelerate them. (*b*) On average, the kinetic energy does not change over time because the average final velocity (drift velocity) does not depend on time. (*c*) Because the electric field does work on the electrons but their kinetic energy does not increase over time, some other form of energy in the system must increase. (As you'll see shortly, the increase takes the form of thermal energy.)

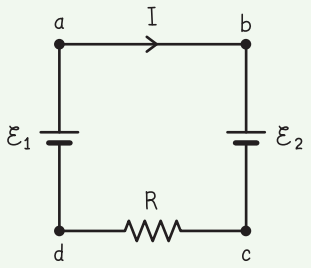
31.12 (*a*) Greater vibrations cause the ions to move around within a greater volume, increasing the probability of ion-electron collisions and decreasing the average time interval between collisions. Because conductivity is proportional to this time interval and resistance is inversely proportional to conductivity, decreasing the interval causes the resistance to increase. (*b*) A current causes the metal to heat up because the work done on the electrons by the applied electric field gets converted to thermal energy of the lattice ions. (*c*) See Figure S31.12 on the next page. At low current, the heating is negligible. Because metals are ohmic at constant temperature, the curve is a line with slope $1/R$ at low current. At high current, which causes the metal to heat up, the resistance increases as the current increases, and so the line curves downward.

Figure S31.12



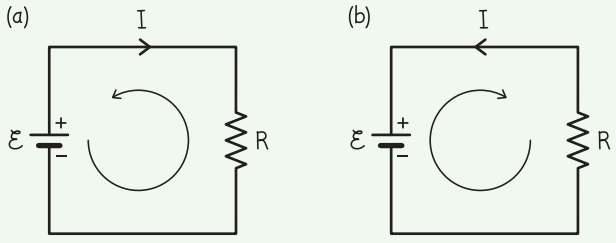
31.13 No. Because the positive terminals of the two batteries are connected to each other with a wire, the potential is the same at points a and b (Figure S31.13). So, if $\mathscr{E}_1 < \mathscr{E}_2$, then the potential at d must be greater than the potential at c. Because the direction of current is from high to low potential, the current is counterclockwise though the circuit, in the direction opposite the reference direction for the current indicated in the diagram.

Figure S31.13



31.14 (*a*) See Figure S31.14*a*. If we go counterclockwise from the top, the first circuit element is the battery; because we are traveling from the positive to the negative terminal, the potential difference is $-\mathscr{E}$. Next is the resistor; because we are traveling in the opposite direction as the reference direction for the current, the potential difference is $+IR$. Substituting these values in Eq. 31.21, we get $-\mathscr{E} + IR = 0$ or $I = \mathscr{E}/R$, which is identical to the result we obtained in Eq. 31.24. We should get the same sign for the current because we have chosen the same reference direction for the current as in Figure 31.34. (*b*) See Figure S31.14*b*. If we go clockwise from the bottom, the first circuit element is the battery; because we are traveling from the negative to the positive terminal, the potential difference is again $\mathscr{E}$. Next is the resistor; because we are traveling in the opposite direction as the reference direction for the current, the potential difference is IR. Substituting these values in Eq. 31.21, we get $\mathscr{E} + IR = 0$ or $I = -\mathscr{E}/R$, which is opposite in sign to the result we obtained in Eq. 31.24. The negative sign means that the current direction is opposite the chosen reference direction for the current. That is, the current direction is not counterclockwise, but clockwise, in agreement with our earlier analysis.

Figure S31.14



31.15 (*a*) The potential difference is now positive because you are moving closer to the positive terminal of the battery. The magnitudes of the current and resistance are the same, however, so the magnitude of the potential difference must be the same. Thus, $V_{cb} = IR_1$. (*b*) From the solution I obtained in Exercise 31.8, $I = (9\text{ V})/(10\ \Omega + 5\ \Omega) = 0.6$ A. Therefore, $V_{cb} = (0.6\text{ A})(10\ \Omega) = 6$ V. This result makes sense because the resistance of R_1 is two-thirds of the resistance of the combination of R_1 and R_2, which means you expect the potential difference across R_1 to be two-thirds of 9 V, the potential difference across the combination.

31.16 (*a*) From Eq. 31.32, $R_{eq} = [1/(3\ \Omega) + 1/(10\ \Omega) + 1/(5\ \Omega)]^{-1} = 1.6\ \Omega$. (*b*) From Eq. 31.31, $I = (9\text{ V})/(1.6\ \Omega) = 5.7$ A.

31.17 Now the resistance in the path from a to d is smaller than it is when $R_{var} = 12\ \Omega$. More charge carriers now flow from a to d, so a small number of carriers flow from b to a through the bulb to increase the current from a to d.

31.18 Because 1 A = 1 C/s and 1 V = 1 J/C = 1 A · Ω,

$$1\text{ A}\cdot\text{V} = 1\frac{\text{C}}{\text{s}}\cdot\frac{\text{J}}{\text{C}} = 1\text{ J/s} = 1\text{ W}$$

$$1\text{ A}^2\cdot\Omega = 1\text{ A}\cdot\text{A}\cdot\Omega = 1\text{ A}\cdot\text{V} = 1\text{ W}.$$

31.19 (*a*) If I travel counterclockwise around the circuit starting at *a*, the loop rule yields $6.0\text{ V} + I(0.25\ \Omega) + I(0.25\ \Omega) - 9.0\text{ V} = 0$. Solving this expression for the current again yields $I = 6.0$ A, so the answer remains unchanged (as it should because nothing has physically changed). (*b*) Because the current is the same in both batteries and because they have the same internal resistance, the rate at which energy is dissipated is also the same: $P = 9.0$ W.

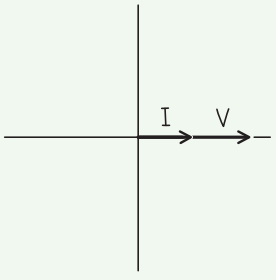
Chapter 32

32.1 (*a*) Electric potential energy in the electric field of the capacitor. (*b*) As the capacitor discharges, that energy is converted to magnetic energy stored in the magnetic field in the inductor. (*c*) Magnetic energy in the inductor.

32.2 (*a*) Yes (see Eq. 31.43). (*b*) Energy is supplied by the AC source, and so as long as the AC source delivers energy at the rate at which it is dissipated by the resistor, the amplitudes of the current and potential difference oscillations remain constant. (Because the power is proportional to the current *squared*, the changing sign of the current does *not* imply that sometimes the resistor delivers energy to the AC source. We discuss this point in Section 32.8.)

32.3 See Figure S32.3. Because the current and potential difference are in phase with each other, the phasors V_R and I overlap as in Figure 32.13. Because v_R and i are both zero at $t = 0$ and increase as time goes on, the phasors must lie on the horizontal axis.

Figure S32.3



32.4 Figure 32.13: zero for both v_R and i. Figure 32.20: zero for i and $\pi/2$ for v_L.

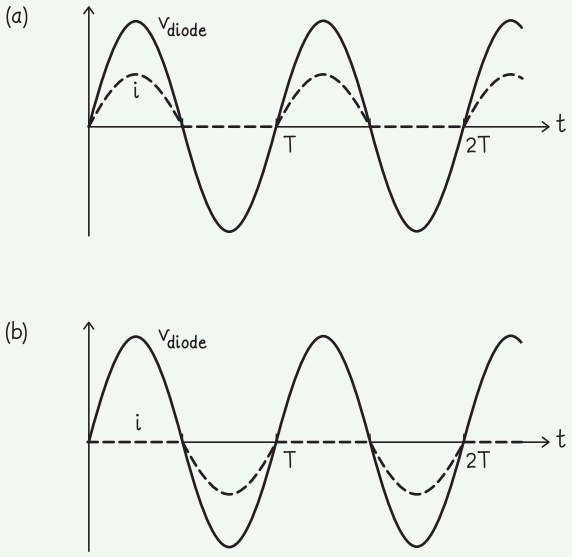
32.5 A piece of doped silicon, whether *n*-type or *p*-type, is electrically neutral because the dopant atoms have the same number of protons as electrons. Although in a doped semiconductor there is either

a surplus or a deficit of electrons relative to the number of electrons in a perfect silicon lattice, there is also an equal surplus or deficit of protons in the nuclei of the dopant atoms.

32.6 Holes travel from left to right (but only in the *p*-type region; in the *n*-type region they recombine with free electrons); simultaneously, electrons travel from right to left.

32.7 The current in the diode is in one direction only. If you assume the diode is attached in the circuit so that it conducts current only when $v_{diode} > 0$, your sketch looks like Figure S32.7*a*. If you assume the diode is attached so that it conducts current only when $v_{diode} < 0$, your sketch looks like Figure S32.7*b*.

Figure S32.7



32.8 According to the junction rule, you must have $I_b + I_c = I_e$.

32.9 The AND gate is the one in Figure 32.39*a*. Both field-effect transistors function as switches and are closed (conducting) when their gates are positively charged. If the potential at A is positive with respect to ground, transistor 1 is conducting. If the potential at B is positive with respect to ground, transistor 2 is conducting. Because the transistors are connected in series, both must be conducting in order for the output potential to be different from ground, and so both A and B must be positive to obtain a nonzero output.

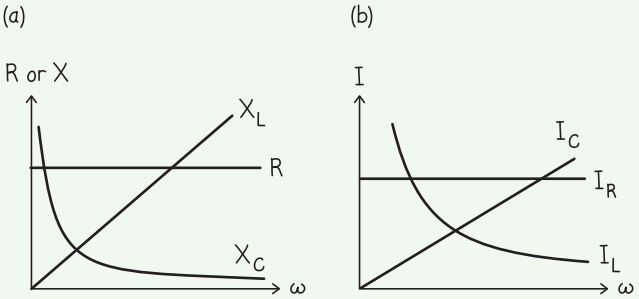
The OR gate is the one in Figure 32.39*b*. Again, the two transistors function as switches, but this time they are connected in parallel. Therefore only one of them needs to be conducting in order for the output potential to be nonzero, and so if either A or B (or both) is positive we obtain a nonzero output.

32.10 (*a*) Higher. (*b*) Because R is always positive, Eq. 32.2 tells you that v_R must be positive when i is positive. According to Eq. 32.3, $\mathscr{E}$ is positive, too. Given that v_R is positive and that $v_a > v_b$, you must have $v_R = v_a - v_b$. (*c*) Because the current is now negative, v_R and $\mathscr{E}$ must be negative, too. (Indeed, with the current now running counterclockwise, v_a is now smaller than v_b.) The potential difference across the resistor is still given by $v_R = v_a - v_b$.

32.11 The current in the circuit. The velocity of a simple harmonic oscillator is equal to the time derivative of its position. Because the potential difference v_C is proportional to the charge on the upper capacitor plate, the rate at which v_C changes—that is, the time derivative of v_C—is proportional to the rate at which the amount of charge on the plate changes, dq/dt. This quantity is the current.

32.12 See Figure S32.12. The two horizontal lines mean that the resistance and the current in the resistor do not depend on angular frequency. The hyperbolic X_C curve shows that the capacitive reactance is inversely proportional to ω (Eq. 32.14). Consequently the amplitude of the current in the capacitor is directly proportional to ω, which means the I_C curve has a constant slope. The straight, positive-slope X_L curve means that the inductive reactance is directly proportional to ω (Eq. 32.26); consequently the current in the inductor is inversely proportional to ω, as shown by the hyperbolic I_L curve.

Figure S32.12



32.13 No. You cannot use the phasor method because the two phasors for the potential differences do not rotate as a unit with a constant phase difference.

32.14 We analyze this circuit as we did in Example 32.8, except that now $V_{out} = V_C$ instead of $V_{out} = V_R$. To take the high-ω and low-ω limits of V_{out}, we rewrite Eq. 32.36 as

$$V_{out} = V_C = \frac{\mathscr{E}_{max}}{\sqrt{\omega^2 R^2 C^2 + 1}}.$$

The low-ω limit of V_{out} is then $V_{out} = \mathscr{E}_{max}$, while the high-$\omega$ limit is $V_{out} = \mathscr{E}_{max}/\omega RC$, which approaches zero for large ω. This circuit therefore passes low-frequency signals essentially unchanged and attenuates high-frequency signals.

32.15 (*a*) The maximum potential differences can be found from the maximum current (using Eqs. 32.6, 32.13, and 32.24). Rounding off to two significant digits, you have $V_R = IR = 1.6 \times 10^2$ V, $V_L = I\omega L = 3.1 \times 10^2$ V, and $V_C = I/\omega C = 3.1 \times 10^2$ V.

(*b*) No. $(V_R + V_L + V_C) \gg \mathscr{E}_{max}$ because the potential differences are not in phase and are never simultaneously at their maxima.

32.16 Changing L or C while keeping R constant changes the resonant angular frequency (unless both L and C are changed in such a way that their product remains constant) and can change the shape of the $I(\omega)$ curve. Increasing C while keeping L fixed broadens the curve, and increasing L while keeping C fixed sharpens the curve. Because it depends only on R and on the applied emf, the amplitude of the current at resonance is not affected by the values of L or C.

32.17 Above, because when $V_L > V_C$, the inductor dominates the reactance.

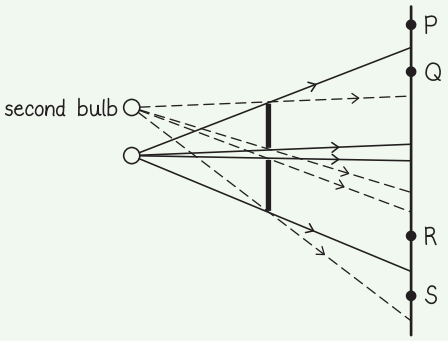
32.18 We substitute the answers from Example 32.10 into Eq. 32.52:

$$P_{av} = I_{rms}^2 R = \frac{I^2}{(\sqrt{2})^2} R = \frac{(3.2\text{ A})^2(50\ \Omega)}{2} = 2.6 \times 10^2\text{ W}.$$

Chapter 33

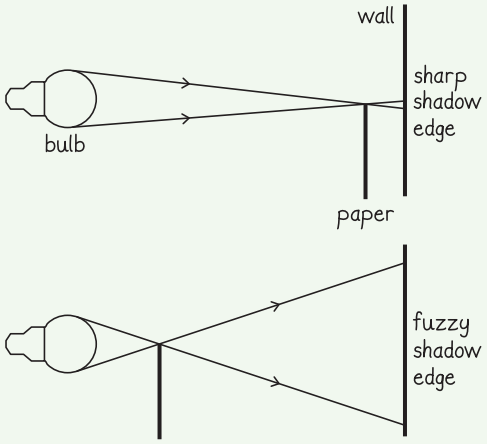
33.1 (*a*) See Figure S33.1. In considering whether the brightness of any location on the screen has changed, note that the distribution of light from the first bulb has not changed, which means the brightness of a particular location changes only if additional light is cast on that location by the second bulb. If the second bulb does not cast light on a particular location, the brightness does not change. (The fact that a given location would be shadowed if it were illuminated by only the second bulb does not decrease the brightness of the light from the first bulb.) The brightness of the spot created by the first bulb does not change because no light from the second bulb strikes this spot. (*b*) Locations P and Q are now brighter than before because some light from the second bulb strikes them. Locations R and S are unaffected because no light from the second bulb reaches them.

Figure S33.1



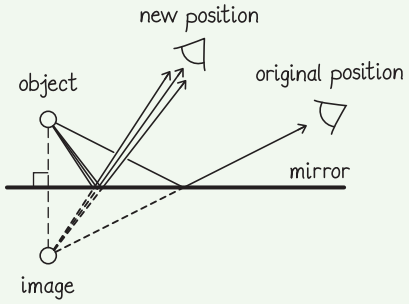
33.2 See Figure S33.2. The shadow edges become sharper as you move the paper farther from the bulb and fuzzier as you move the paper closer to the bulb. This happens because the bulb is not a point source of light. Rays from different parts of the bulb's surface pass by the edge of the paper at slightly different angles. When the paper is farther from the bulb, the difference in these angles is less and so the edge of the shadow is sharper. When the paper is close to the bulb, the angular size of the filament (as seen from the edge) is greater, which makes the edge of the shadow fuzzier.

Figure S33.2



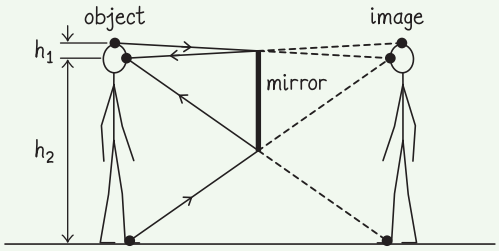
33.3 No. Figure S33.3 shows the reflected rays that reach the observer in the two locations. Because the angle of incidence always equals the angle of reflection when the reflecting surface is smooth, any ray from the object that reflects anywhere on the mirror can be traced back through the mirror to a location directly behind the object. Moving the observer changes only which subset of the reflected rays the observer sees and not where the rays appear to come from.

Figure S33.3



33.4 No. Figure S33.4 shows the rays that reflect into the person's eye from the highest and lowest points on his body. These rays strike the mirror at a height midway between where they originate on the person and the person's eye. So the mirror needs to extend only from halfway between the person's eye and the highest point on the person's body to halfway between the person's eye and the lowest point on the person's body. The ray from the highest point strikes the mirror at a height midway between the height h_1 in Figure S33.4 at which the ray originates and the height at which it strikes the eye, and the same is true for the ray from the lowest point. So the top of the mirror can be at a height that is half the height h_1 above the height of the eye, and the bottom of the mirror can be at a height that is half the height h_2 below the height of the eye. Because $h_1 + h_2$ is the person's height, this means the mirror needs to be only half this height. (However, the mirror must be positioned at the right height.)

Figure S33.4

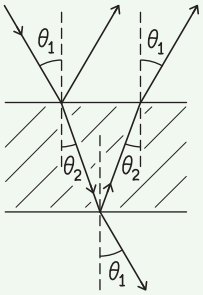


33.5 The wavefronts arrive at a given location in the glass at the same frequency that they arrive at any location in vacuum. However, they travel only two-thirds as fast in the glass as in vacuum. Sequential wavefronts arrive at the glass surface at instants separated by the period T of the wave ($T = 1/f$). If the wave traveled at the same speed in glass as it does in vacuum, then during one period, one wavefront would travel a distance into the glass equal to the vacuum wavelength (400 nm) before the next wavefront arrived at the surface of the glass, and the wavefronts would be separated by 400 nm. However, because the wavefronts travel at only two-thirds the speed of light in vacuum, the spacing between wavefronts in the glass must be two-thirds of the vacuum wavelength, or 267 nm.

33.6 The propagation of the wavefronts is exactly like that shown in Figure 33.16*a* with the direction of propagation reversed. Consequently, because the angle of incidence equals θ_2, the angle of refraction must equal θ_1.

33.7 The angle between the ray reflected from the bottom surface and the normal is θ_2, as shown in Figure S33.7. This ray is now incident at the top surface and refracted as it emerges into the air. This is equivalent to the situation shown in Figure 33.17*b*, in which the ray originates in the glass and is refracted in the air. Therefore, as discussed in Example 33.3, in air the angle this refracted ray makes with the normal is θ_1. Note in Figure S33.7 that this refracted ray is shifted sideways relative to the ray *reflected* from the top surface.

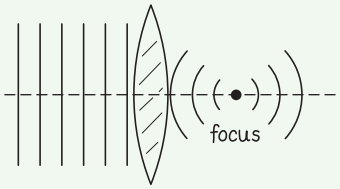
Figure S33.7



33.8 The critical angle corresponds to the angle of incidence for which the angle of refraction is 90°. When entering the medium, to obtain a given angle of refraction, the angle of incidence of a violet ray must be greater than that of a red ray; therefore the critical angle is smaller for a violet ray than for a red ray.

33.9 See Figure S33.9. The wavefronts are perpendicular to the light rays everywhere.

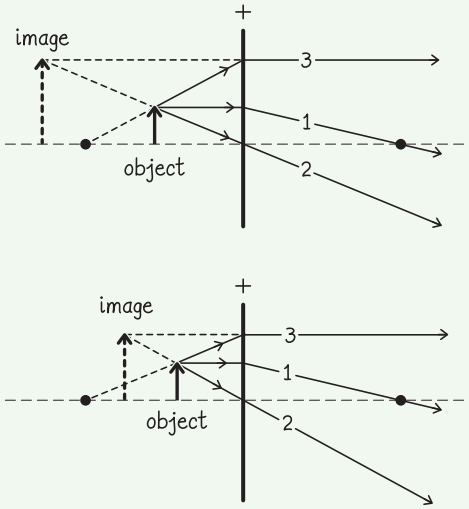
Figure S33.9



33.10 No. The image location can be found with any two of the three principal rays. Because all three intersect at one point, any two of them indicate the point of intersection and therefore the image location. It is, however, useful to draw all three rays in order to check that you have made no mistakes.

33.11 Decrease. See Figure S33.11. As the object moves closer to the lens, ray 3 makes a smaller angle with the lens axis and thus intercepts the lens closer to the axis. Ray 1 remains the same, and ray 2 makes a greater angle with the axis in order to pass through the center of the lens. As a result, the image point—the virtual intersection of the rays—is closer to the axis (as well as the lens), making the image smaller.

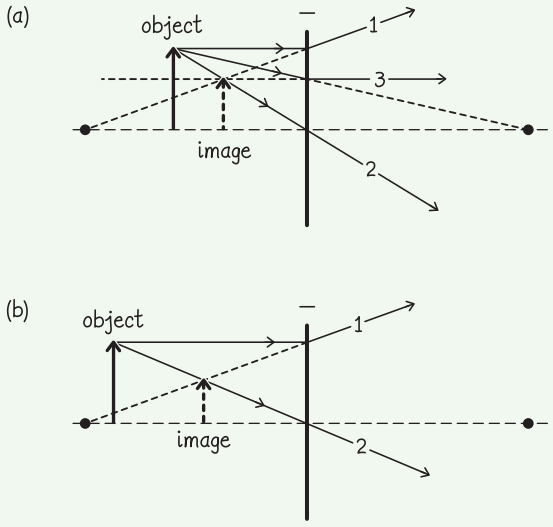
Figure S33.11



33.12 Virtual, because the rays do not actually intersect anywhere on the image. They merely appear to originate from a common point on the virtual image.

33.13 (*a*) See Figure S33.13*a*. You draw ray 3 by first drawing a line connecting the point of interest on the object with the focus on the side of the lens opposite the object. Run ray 3 along this line from the object to the lens surface; then change the ray direction so that it emerges parallel to the lens axis. (*b*) There is no position for which the image is larger than the object. To see why, all you need to consider is principal rays 1 and 2. When the object is moved, the direction of ray 1 does not change—only ray 2 changes direction. When the object is moved to the left, as in Figure S33.13*b*, ray 2 now intersects ray 1 closer to the lens axis and the image is smaller. Moving the object right, toward the lens, makes the image larger, but it cannot become larger than the object because, on the left side of the lens where the image forms, ray 1 is always below the tip of the object. As shown in Figure 33.33, placing the object beyond the object-side focus also produces an image that is smaller than the object. (*c*) It is not possible to form a real image with a diverging lens because rays diverge when they go through it.

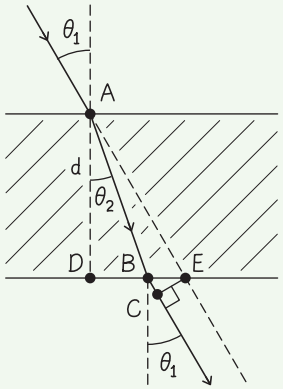
Figure S33.13



33.14 See Figure S33.14. (*a*) You need to determine the distance CE in terms of the angle of incidence θ_1, slab thickness d = AD, and slab index of refraction n_2. Angle CBE = $90° - \theta_1$, and so angle BEC = θ_1. Therefore CE = BE cos θ_1. To determine the distance BE, you can say DB + BE = DE, DE = d tan θ_1, and DB = d tan θ_2.

Substituting the second and third expressions in the first, you obtain d tan θ_2 + BE = d tan θ_1, which can be solved for BE:

Figure S33.14



$$\text{BE} = d(\tan\theta_1 - \tan\theta_2) = d\left(\frac{\sin\theta_1}{\cos\theta_1} - \frac{\sin\theta_2}{\cos\theta_2}\right). \qquad (1)$$

Applying Snel's law (Eq. 33.7) to this case and solving Eq. 33.7 for $\sin\theta_2$ gives you $\sin\theta_2 = n_1 \sin\theta_1/n_2$. Next you can use the identity $\sin^2\theta + \cos^2\theta = 1$ to get an expression for $\cos\theta_2$:

$$\cos\theta_2 = \sqrt{1 - \sin^2\theta_2} = \sqrt{1 - \left(\frac{n_1 \sin\theta_1}{n_2}\right)^2}.$$

You can substitute these expressions for $\sin\theta_2$ and $\cos\theta_2$ into Eq. 1:

$$\text{BE} = d\left[\frac{\sin\theta_1}{\cos\theta_1} - \frac{n_1 \sin\theta_1}{n_2\sqrt{1 - \left(\frac{n_1 \sin\theta_1}{n_2}\right)^2}}\right],$$

which simplifies to

$$\text{BE} = d\sin\theta_1\left(\frac{1}{\cos\theta_1} - \frac{n_1}{\sqrt{n_2^2 - n_1^2\sin^2\theta_1}}\right).$$

From this you can determine the distance CE:

$$\text{CE} = \text{BE}\cos\theta_1 = d\sin\theta_1\left(1 - \frac{n_1\cos\theta_1}{\sqrt{n_2^2 - n_1^2\sin^2\theta_1}}\right).$$

$$(b)\ \text{CE} = (0.010\text{ m})(0.50)\left(1 - \frac{(1)(0.87)}{\sqrt{(1.5)^2 - (1)^2(0.5)^2}}\right)$$

$$= 0.0019\text{ m}.$$

Thus, at an angle of incidence of 30°, in passing through a glass slab that is 10 mm thick (about three times the thickness of a typical windowpane), a light ray is shifted sideways by 2 mm (20% of the thickness of the slab).

33.15 (*a*) f is negative (lens is diverging), so $f = -80$ mm; o is positive (object is on same side as illumination), so $o = 100$ mm. Substituting these values into Eq. 33.16 gives

$$\frac{1}{i} = \frac{1}{-80\text{ mm}} - \frac{1}{100\text{ mm}} = -0.0225\text{ mm}^{-1},$$

and solving for i gives $i = -44$ mm. The value of i is negative, as it should be for a virtual image, and the absolute value of i is less than o, which is consistent with Figure 33.33.

(*b*) Equation 33.17 gives

$$M = \frac{h_i}{h_o} = \frac{-i}{o} = \frac{-(-44\text{ mm})}{100\text{ mm}} = 0.44.$$

The image height is 44% of the object height.

33.16 (*a*) Solving Eq. 33.21 for f gives $f = (0.25\text{ m})/M_\theta$, and substituting this result into Eq. 33.22 gives $d = 4M_\theta$. For $M_\theta = 8$, you have $d = +32$ diopters.

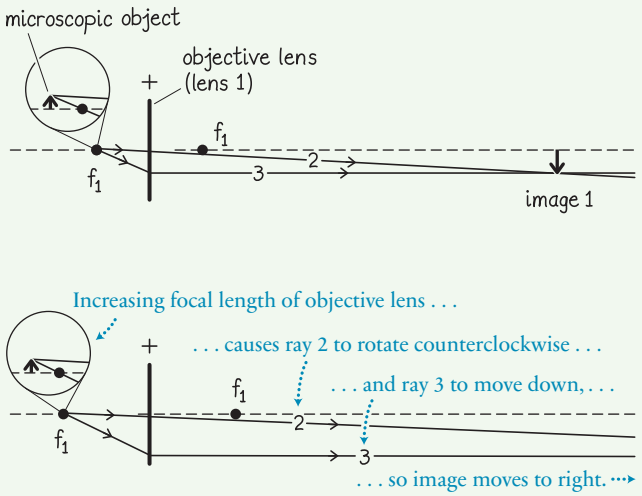
(*b*) $f = (0.25\text{ m})/M_\theta = (0.25\text{ m})/8 = 0.031\text{ m} = 31\text{ mm}$.

33.17 (*a*) The image moves farther from the objective lens. This can be seen either by constructing a ray diagram (Figure S33.17) or by considering the relationship among object distance, image distance, and focal length (Eq. 33.16). Solving Eq. 33.16 for i_1 gives

$$\frac{1}{i_1} = \frac{1}{f_1} - \frac{1}{o_1} = \frac{o_1 - f_1}{o_1 f_1} \text{ or } i_1 = \frac{o_1 f_1}{o_1 - f_1}.$$

Increasing both o_1 and f_1 will increase the numerator of the expression for i_1; keeping the sample "just outside the focal point" of the lens implies that the distance between the sample and the lens, $o_1 - f_1$, is kept at least roughly the same, so increasing the numerator while keeping the denominator constant increases i_1. (Practical limitations on the construction of lenses actually allow the distance $o_1 - f_1$ to be smaller for shorter-focal-length lenses, so in fact, the denominator also increases as the focal length increases. However, this last effect is not a consequence of the simple thin-lens treatment but of ways that lenses are not ideal that are beyond the scope of this text.)

Figure S33.17



(*b*) Shorter focal lengths give a more compact microscope, because the first image must fall close to the focal point of the eyepiece lens, and thus the distance between the two lenses is roughly $i_1 + f_2$. From part *a* we know that i_1 depends on f_1, and so decreasing f_1 decreases i_1.

33.18 (*a*) Using the result of Example 33.9 gives you $f_1 = M_\theta f_2 = (22)(0.0400\text{ m}) = 0.88$ m.

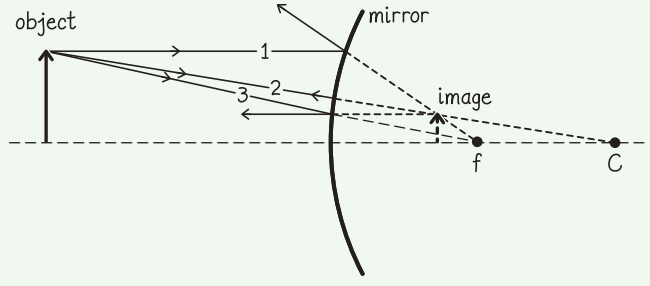
(*b*) The length of the telescope is roughly the sum of the focal lengths of the two lenses (see Figure 33.50): 0.88 m + 0.04 m = 0.92 m. Because the eyepiece lens is by design a short-focal-length lens, the length of the telescope is determined primarily by the focal length of the objective lens.

33.19 (*a*) The ray diagram in Figure S33.19 shows that the image is behind the mirror and virtual. You calculate the image distance with Eq. 33.24. The focal length is half the radius of curvature, C = 1.0 m, and is negative because it is a virtual focus:

$$i = \left(\frac{1}{f} - \frac{1}{o}\right)^{-1} = \left(\frac{1}{-0.50} - \frac{1}{1.0}\right)^{-1} = -0.33\text{ m}.$$

The negative value for i tells you that the image is located 0.33 m behind the mirror.

Figure S33.19



(*b*) $h_i = -\dfrac{h_o i}{o} = -\dfrac{(0.30\text{ m})(-0.33\text{ m})}{(1.0\text{ m})} = 0.10\text{ m}.$

The image is smaller than the object, and the positive value for h tells you that the image is upright.

33.20 The only change to the physics behind the derivation of the lensmaker's formula is that Eq. 33.25 becomes $n_1 \sin\theta_i = n_2 \sin\theta_r$. In the paraxial approximation, this result can be written as $(n_1/n_2)\theta_i = \theta_r$. Comparing this with Eq. 33.26, you see that the effect of submerging the lens is to substitute (n_1/n_2) for n. Once you make this change in the remainder of the derivation, the lensmaker's formula becomes $\frac{1}{f} = \left(\frac{n_1}{n_2} - 1\right)\left(\frac{1}{R_1} + \frac{1}{R_2}\right)$.

Chapter 34

34.1 Once the fronts reach P, the intensity at P is the same as the intensity at any point on any of the wavefronts. For any point R above or below the ray through P, you can locate a corresponding point R′ for which the distance R′P is exactly half a wavelength greater than the distance RP. Those points therefore do not contribute to the intensity at P. Only the point exactly on the ray (dashed line) contributes.

34.2 If you identify pairs of points that radiate in phase at P, then, for each pair, there exists another pair that is exactly 180° out of phase with the first pair and thus cancels the radiation from the first pair.

34.3 (*a*) The spacing between fringes decreases as the separation between the slits increases because $\sin\theta$ is proportional to $1/d$. (*b*) The spacing between fringes increases as the wavelength increases because $\sin\theta$ is proportional to λ.

34.4 No. Two of the waves cancel each other perfectly, but the third wave is not canceled. The intensity observed on the screen, at angles given by $d \sin\theta = \pm(n - \frac{1}{2})\lambda$, is the intensity of one of the three waves.

34.5 Fringe brightness is determined by the intensity (power/area) that strikes the screen. Increasing the number of slits increases the amount of power that travels through the slits to the screen. In addition, increasing the number of slits causes the fringes to sharpen, with the result that the power that reaches the screen is concentrated into narrower fringes, decreasing the area and thus increasing the brightness further.

34.6 (*a*) As you found in Checkpoint 34.3, the angle of the fringes increases with wavelength, and so the angle is less for violet light than for red. White light is thus dispersed into a rainbow, with the violet end of the spectrum at smaller angles than the red. (*b*) Each rainbow corresponds to the fringes of a particular order for all colors, or said another way, each order produces its own rainbow.

34.7 (*a*) The spacing between fringes in a two-slit interference pattern increases with increasing wavelength and with decreasing distance between slits (see Checkpoint 34.3). The same is true for diffraction gratings: The separation distance between adjacent fringes is increased by decreasing the separation distance between adjacent slits. Therefore the distance between adjacent slits should be decreased if your aim is to increase the diffraction grating's ability to separate close wavelengths. (*b*) No, as long as the slits are very narrow. The fringe spacing is determined entirely by the separation distance between adjacent slits. (Slit width does, however, affect the brightness of the pattern by determining how much light gets through the diffraction grating.)

34.8 Yes. The diffracted x rays interfere constructively at all angles for which the difference in path length is either zero (as it is for $\theta' = \theta$) or an integer multiple of λ. The diffracted beam at $\theta' = \theta$ corresponds to the zeroth-order maximum; other beams correspond to higher orders and are much weaker.

34.9 From Figure 34.20*b* you can see that $2\theta + 2\alpha = \pi$, so $\theta = -\alpha + \frac{\pi}{2}$. Using trigonometry, I get $\cos\theta = \cos(-\alpha + \frac{\pi}{2}) = -\sin(-\alpha) = \sin\alpha$. Substituting this into the Bragg condition yields $2d \sin\alpha = m\lambda$.

34.10 Crystal B. Spots closer together correspond to greater distances between crystal planes and hence greater atomic spacing.

34.11 It decreases, because the spacing between spots is proportional to the wavelength.

34.12 If the electrons travel more slowly, their wavelength increases, and the diffraction spots spread farther apart.

34.13 You could perform the experiment shown in Figure 34.6 with a light source so weak that the light "particles" pass through the pair of slits only one at a time. If light indeed has particle properties, a source this weak would give you a pattern that initially is like the one in Figure 34.25*a*, showing where individual particles of light hit the screen.

34.14 The maximum probability is at the locations of greatest intensity in the interference pattern; these locations are at angles θ for which $\sin\theta$ is a multiple of λ/d.

34.15 Although the maximum time-averaged intensity is greater than the sum of the intensities of the original two beams, averaging this time-averaged intensity over the entire area filled by the bright and dark bands of the interference pattern gives just the sum of the intensities of the original two beams, $2S_{0,av}$ (because the average of the cosine squared is one-half).

34.16 No, because the higher-order bright fringes exist only as long as $m\lambda/d < 1$. For values of m such that $m\lambda/d > 1$, there are no bright fringes because these correspond to $\sin\theta_m > 1$, for which there is no angle θ_m.

34.17 The oil forms a thin film that causes thin-film interference. The thickness of the oil layer determines for which wavelength constructive interference occurs. If the thickness of the film varies spatially, different wavelengths interfere constructively at different locations.

34.18 (*a*) If $a < \lambda$, there are no dark fringes because that would require $\sin\theta > 1$. This indicates that no minimum exists, and the light coming through an aperture with a diameter less than the wavelength acts as a true point source as assumed in our original discussion of multiple-slit interference. When $a < \lambda$, the slit acts as a point source with the wavelets spreading out spherically from the slit. (*b*) For $a \gg \lambda$, the first dark fringe occurs very close to $\theta = 0$. This means there is essentially no diffraction, and the beam just propagates straight ahead. This illustrates that single-slit diffraction is observed primarily with slits from a few wavelengths wide to tens of wavelengths wide.

34.19 Resolution is determined by wavelength and the ratio f/d: $y_r = 1.22\,\lambda f/d$. For a given wavelength, the greater f/d is, the larger the Airy disk. For the three lenses, f/d is (*i*) 1.3, (*ii*) 1.5, and (*iii*) 1.1. (*a*) The highest resolution is obtained with the smallest Airy disk, produced by lens *iii*. (*b*) The lowest resolution is obtained with the largest Airy disk, produced by lens *ii*. (The wavelength determines the exact size of the Airy disk; the comparisons made here assume the same wavelength for all three lenses.)

34.20 (*a*) The stopping potential difference V_{stop} is proportional to the frequency (Figure 34.48*b*) of the incident photons, and the maximum kinetic energy K_{max} of the ejected electrons is proportional to V_{stop} (Eq. 34.34). Therefore, K_{max} must be proportional to the frequency of the incident photons. (*b*) The minimum frequency the light can have and be able to eject electrons. Lower-frequency light does not eject electrons.

34.21 (*a*) The photon slows down because the speed of light is less in a medium that has an index of refraction of 1.5 than in air (index of refraction 1). (*b*) The frequency remains unchanged, as discussed in Chapter 33. (*c*) The wavelength decreases because the wavefronts travel less far in a given time interval. (*d*) The photon's energy does not change because the medium does not take away any energy. This is why photon energy must be expressed in terms of frequency rather than wavelength, because neither frequency nor energy depends on the medium, whereas wavelength does.

This page intentionally left blank

Credits

COVER PHOTO shihina/Shutterstock

CHAPTER 1 **P.19** Martin Prague/Shutterstock; **p. 23 left** Eric Schrader/Fundamental Photographs; **right** Leonardo DaVinci/Public Domain; **p. 25 left** Wilson Ho, University of California, Irvine; **right, from top to bottom** Wilson Ho, University of California, Irvine, José Antonio Peña/Science Source, Torsten Wittmann/Science Source, John Woodcock/Getty Images, NASA, NASA/GSFC/SDO, Roberto Rizzo/123RF, NASA, NASA/ESA, NASA; **p. 27 from top to bottom** Robert Lucian Crusitu/Shutterstock, Waddell Images/Shutterstock, Fedor Selivinov/123RF, Bombaert /123RF; **p. 28** dziewul/123RF.

CHAPTER 2 **P. 53** blanscape/Shutterstock; **p. 54** both Eric Mazur **p. 56** Eric Mazur **p. 58** Eric Mazur; **p. 60** both Eric Mazur; **p. 66** Eric Mazur; **p. 68** Eric Mazur.

CHAPTER 3 **P. 86** Avigator Fortuner/Shutterstock; **p. 89** Igor Kovalchuk/Shutterstock; **p. 89** David T. Sandwell and Walter H. F. Smith/NOAA.

CHAPTER 4 **P. 116** Dabarti CGI/Shutterstock; **p. 117** PASCO Scientific; **p. 118** Eric Mazur; **p. 120** Accent Alaska/Alamy Stock Photo; **p. 139** Eric Mazur.

CHAPTER 5 **P. 149** Xurzon/Getty Images; **p. 153 left** Evgeny Murtola/Shutterstock; **right** Stephen Rudolph/Shutterstock.

CHAPTER 6 **p. 176** Sarunas Vaitkus/Shutterstock; **p. 191** Denis Doyle/Bloomberg/Getty Images.

CHAPTER 7 **P. 211** image BROKER/Alamy Stock Photo; **p. 215** Ted Kinsman/Science Source; **p. 225** Can Inellioglu/Shutterstock; **p. 226** NASA, **inset** Steve Heap/Shutterstock.

CHAPTER 8 **P. 246** nikkytok/Shutterstock.

CHAPTER 9 **P. 282** Sergey Borisov_88/Shutterstock.

CHAPTER 10 **P. 314** Arthur Bargan/Alamy Stock Photo; **p. 315** Eric Mazur; **p. 317** NASA Earth Observatory; **p. 328** Eric Mazur.

CHAPTER 11 **P. 351** shihina/Shutterstock; **p. 366** Eric Schrader/Pearson Education; **p. 371** Toby Melville/Reuters.

CHAPTER 12 **P. 388** Jari Hindstroem/Shutterstock; **p. 393** Richard Megna/Fundamental Photographs.

CHAPTER 13 **P. 425** Joggie Botma/Shutterstock; **p. 434** NASA; **p. 437 top** NASA Earth Observatory; bottom NASA; **p. 439** baranozdemir/Getty Images.

CHAPTER 14 **P. 461** Erich Lessing/Art Resource.

CHAPTER 15 **P. 504** Westend61 GmbH/Alamy Stock Photo; **p. 510** Klaus Sailer/Shutterstock.

CHAPTER 16 **P. 539** EyeEm/Alamy Stock Photo; **p. 540** AFP/Getty Images.

CHAPTER 17 **P. 580** tasha0102/Shutterstock; **p. 587** Nokia Corporation; **p. 769** Berenice Abbott/Science Source; **p. 594** Mikael Damkier/Shutterstock; **p. 609** Omikron/Science Source.

CHAPTER 18 **P. 616** Dmitry Naumov/Shutterstock; **p. 624** culture-images GmbH/Alamy Stock Photo; **p. 625 left** Omikron/Science Source; **right** Omikron/Science Source; **p. 629** NASA; **p. 630** Eric Schrader/Pearson Education; **p. 631** Andrei Kuzmik/Shutterstock; **p. 640** xtrekx/Shutterstock.

CHAPTER 19 **P. 664** Evucis/Shutterstock.

CHAPTER 20 **P. 701** hermione13/123RF.

CHAPTER 21 **P. 742** Nick Harrison/Alamy Stock Photo; **p. 758** Colin Salmon/Getty Images.

CHAPTER 22 **P. 783** kyslynskahal/Shutterstock; **p. 784** All Canada Photos/Alamy Stock Photo **p. 785** Eric Schrader/Pearson Education; **p. 791** Paul A. Souders/Corbis/VCG/Getty Images; **p. 794** GIPhotoStock/Science Source.

CHAPTER 23 **P. 813** Toppy Baker/Shutterstock; **p. 817** John W. Belcher; **p. 819** John W. Belcher.

CHAPTER 24 **P. 845** Wichai Cheva Photography/Shutterstock; **p. 847b** John W. Belcher.

CHAPTER 25 **P. 878** Chanchai phetdikhai/Shutterstock; **p. 883** USGS.

CHAPTER 26 **P. 908** Aksana Tsishyna/Shutterstock; **p. 911** James Kirkikis/Shutterstock; **p. 923b** yurazaga/Shutterstock.

CHAPTER 27 **P. 941** Alchemy/Alamy Stock Photo; **p. 942a** Eric Mazur; **p. 942** Alex Kosev/Shutterstock; **p. 946** Cordelia Molloy/Science Source.

CHAPTER 28 **P. 976** ArtWell/Shutterstock; **p. 991b** Yon Marsh/Alamy Stock Photo.

CHAPTER 29 **P. 1011** Monika Wisniewska/Shutterstock.

CHAPTER 30 **P. 1042** Kenny Tong/Shutterstock.

CHAPTER 31 **P. 1079** fStop Images GmbH/Alamy Stock Photo.

CHAPTER 32 **P. 1121** Elena11/Shutterstock.

CHAPTER 33 **P. 1164** Erich Lessing/ArtResource; **p. 1172** Steve Hamblin/Alamy Stock Photo; **p. 1172b** Jan Mika/Shutterstock.

CHAPTER 34 **P. 1207** Victor Tyakht/Alamy Stock Photo; **p. 1208a** left Richard Megna/Fundamental Photographs; **p. 1208b** GualtieroBoffi/Shutterstock; **p. 1210** Ken Kay/Fundamental Photographs; **p. 1211** Eric Mazur; **p. 1213** Eric Mazur; **p. 1214** caimacanul/Shutterstock; **p. 1217** Science History Images/Alamy Stock Photo; **p. 1217** Wolfgang Rueckner; **p. 1219** Eric Mazur; **p. 1220c** Wolfgang Rueckner; **p. 1221** Eric Mazur; **p. 1225** Eric Mazur; **p. 1227** Anyka/Fotolia; **p. 1230** Eric Mazur; **p. 1231** GIP Photo Stock/Science Source.

Index

A

Absolute temperature, 686
Absolute time, 26
Absolute zero, 708
Absorbed light, 1166–1169
Acceleration, 87–115
  average, 87, 96, 104
  of car, 92, 98–99, 104
  centripetal, 427–428, 434
  of closed system, 231–232
  collision and, avoiding, 99
  constant, motion with, 90, 96–99
  definition of, 87
  direction of, 87
  displacement and, 89–90, 105–106
  due to gravity, 88–90, 99–100, 102
  in electric field, 820, 822
  in electric potential energy, 879–880
  free-fall equations, 99–102
  in Galilean transformation equations, 190
  on inclined planes, 102–103
  instantaneous, 103–106, 318–319
  interactions and, 228
  magnitude of, 87, 88
  motion diagrams, 92–94
  nonzero, 92, 131, 822
  of object in free fall, 89–91
  projectile motion and, 90–92
  in rolling motion, 406–407
  symbol for, 87
  translational equilibrium and, 251–252
  unit of, 87
  velocity and, changes in, 87–88
  x components of
    average acceleration and, 87, 96, 104
    constant acceleration and, 96
    curvature of *x(t)* curve related to, 88
    due to gravity, 90, 100
    inclined planes and, 102–103
    instantaneous acceleration and, 104–105
    in projectile motion, 92
Acceleration-*versus*-time graphs, 131
AC circuits, 1124–1128
  power in, 1150–1152
  with two resistors, 1137
AC source, 1124
Action at a distance, 814, 820
Addition
  significant digits in, 37
  of vectors, 69, 317
Adhesive forces, 632
Adiabatic processes, 703, 717
Adiabats, 717
Airy, Sir George, 1231
Airy disk, 1231–1233
Alternating currents, 1122–1124
Altimeter, 636
Ammeter, 1101, 1151
Ampère, Marie, 987
Ampère's law, 987–990
Ampèrian paths, 983–985
Amplitude, 1136
  of waves in one dimension, 559, 569
  of waves in two and three dimensions, 583, 589, 592
AND gate, 1134
Andromeda Galaxy, 463
Angle of incidence, 1166–1173, 1180–1182
Angle of reflection, 1166–1167, 1169, 1171, 1173
Angle of refraction, 1170–1172, 1180–1181
Angular coordinate, 326
Angular momentum, 367–372
  conservation of, 400–404, 434
  gravity and, 431–434
  torque and, 389–392
Angular momentum law, 402
Antennas, 1049–1053
Antineutrino, 793
Antinodes, 562
Aphelion, 429
Apparently weightless, 437
Apparent weight, 437
Archimedes' principle, 623
Area, 71
Area vector, 858
Assumptions, hidden, 21
Atmospheres, pressure given in, 635
Atmospheric pressure, 635–636, 641–642
Atom, 24
  core of, 1128
  substitute, 1129
  valence electrons, 1128–1130
Atomic potential difference, 891–892
Attractive interactions, 212
Attributes, 222
Audible frequency range, 585
Automobile battery, 918
Average acceleration, 87, 96, 104
Average power, 567
Average speed, 59–62
Average velocity, 59–62
Avogadro's number, 33
Axial vectors, 398

B

Balances, 435
Barometer, 640–641
Base, 1132
Batteries, in charge separation and storage, 917–919
Beat frequency, 601
Beating, 599–601
Bernoulli effect, 627
Bernoulli's equation, 644–647
Beta decay, 792–793
Bias potential difference (or bias), 1133–1134
Binomial series expansion, 827
Blood flow, 650
Boltzmann constant, 686
Boulder-Earth system, 216
Boundary effects, 550–554
Bragg angle, 1215–1216
Bragg condition, 1216
Brahe, Tycho, 433
Brayton cycle, 757–759, 770–772
Breakdown threshold, 913–914
Broken chemical bonds, 221
Brownian motion, 665–666
Bulk stress, 617, 618, 619
Buoyancy, fluids and, 622–624

C

Calculus, to determine displacement, 106
Calibrating graphs, 55–56
Candela, 33
Capacitance, 921–924
Capacitors, 911–915
  in charge separation and storage, 911–915
  coaxial, 923–924
  definition of, 911
  with dielectrics, 930–931
  fully charged, 912
  isolated, 912
  magnetic field in, 1057–1058
  metal-slab, 914–915
  parallel-plate, 911–914, 922
  power of, 1070–1071
  spherical, 924
Capillary rise, 633
Capillary tube, 632–633
Carnot, Sadi, 757
Carnot cycle, 757–759, 768–770
Cartesian coordinate system, 317, 326, 898
Causal relationship, 26
Cavendish, Henry, 442–443
Celestial mechanics, 446–451
Cell potential difference, 918
Celsius scale, 708, 709
Center of mass, 193–197
  importance of, 196
  position of, 195–196
  reference frames and, 193–194
  velocity and, 195
Center-of-mass velocity, 195
Central electric force, 799
Centripetal acceleration, 427–428, 434
Change, time and, 26
Chapter summary, overview of, 30
Charge carriers, 785
  force exerted by distributions of, 800–803
  mobility of, 788–793
Charged conducting objects, 854–856
Charge distribution, continuous
  electric field of, 897–899
  electrostatic potentials of, 895–897
Charged particles
  in cavity, 855–856
  in electric field, 824–825
  in electric fields, 958–963, 1046–1049
  in magnetic fields, 996–999
  in motion, magnetic fields of, 976–1010
    Ampère's law, 987–990
    Ampèrian paths, 983–985

Charged particles (*continued*)
current loops and spin magnetism, 978–980
magnetic dipole moment and torque, 980–983
magnetic field of moving charged particle, 996–999
magnetic fields due to currents, 993–996
solenoids, 990–991
source of magnetic field, 977–978
toroids, 991–992
trajectory, 820–821
Charge flow, magnetism and, 946–949
Charge polarization, 793–794
Charge-separating devices, 918
Charge separation and storage, 909–940
batteries in, 917–919
capacitance and, 921–924
capacitors and, 911–915
dielectric constant and, 928–931
dielectrics and, 915–917
Gauss's law in, 931–933
in electric field energy, 925–928
emf and, 926
in Van de Graaff generators, 910–911, 913, 917, 927
in voltic cells, 917–919
Charging by induction, 794
Checkpoints, overview of, 30
Chemical bonds, broken, 221
Chemical energy, 156, 219, 221
Circuit diagram, 1082
Circuit element
definition of, 1080
resistance of, 1084
Circuits. *See* Electric circuits
Circular apertures, 1231–1234
Circular motion, 352–387
angular momentum, 367–372
at constant speed, 352–356
forces and, 356–358
rotational inertia, 358–360
of extended objects, 372–377
rotational kinematics, 362–367
Circular symmetry, 977–978
Classical mechanics, 21
Clocks
desynchronization of twins, 472
synchronizing, 462, 464–465
Closed path, 446
Closed surfaces, 849–851
Closed-system approach, 296
Closed systems, 156–158
acceleration of, 231–232
conservation of energy in, 164–166
definition of, 156
Coaxial capacitor, 923–924
Coefficients
of friction, 336–340
of restitution, 163–164
Coherent deformation, 216–218
Coherent (or mechanical) energy, 217–220, 221, 237
Coherent sources, 589–592
Cohesive forces, 628–629
Collector, 1132, 1235
Collisions
classification of, 150–151
elastic, 150, 160–162
inelastic, 150, 163–164
totally inelastic, 150–155, 163, 164, 166
momentum
conservation of momentum and, 135–140
friction and, 117–122
inertial standard and, 127–128
in isolated systems, 133
momentum changes and, 129–130
in two dimensions, 330–331
Color charge, 226
Commutator, 982
Component vectors, 326–327
Compression, 584
Compressional wave, 584
Compressive stress, 617, 618, 619
Concepts, definition of, 24
Condenser, 759
Conducting objects, charged, 854–856
Conducting plate, 866
Conduction, 790
Conductivity, 1093–1094
Conductors
capacitance of, calculating, 921
electric fields in, 1088–1090
moving, in magnetic fields, 1012–1014
Configuration state, 216
Conic sections, 447–448
Conservation
of charge, principle of, 792
of energy, 164–166
law of, 150, 153
special relativity and, 492–496
laws, 26
energy, 150, 153
for interacting objects, 181–184
momentum, 152, 201, 261
relativity and, 201–202
of momentum
angular, 400–404, 434
in collisions, 129–130
in interactions, 135–140
law of, 152
special relativity and, 488–492
in reference frames, 177
in inertial reference frames, 180–181
in interacting objects, 181–184
relativity and, 201–202
in zero-momentum reference frames, 186
Constant acceleration, motion with, 96–99
Constant velocity, 87, 88, 90, 96, 98, 179
Constrained quasistatic processes, 717
Constraints, on systems, 123–124
Constructive interference, 549
Contact angle, 632
Contact forces, 250, 253
Context-rich (CR) problems, 30
Continuity equation, 645
Continuous charge distribution
electric field of, 897–899
electrostatic potentials of, 895–897
Contraction, length, 484–488
Converging lens, 1173–1177, 1183–1186
Conversion factors, 34–35
Convertible kinetic energy, 198–201
Converting energy, 743–746
Convex surface, 1173
Coordinates, in motion, 316
Coulomb (C), 797–798
Coulomb, Charles, 796, 797
Coulomb's law, 796–800, 801, 824, 827, 846, 860
Counting numbers, 37
Creation
to change value of extensive quantity, 123–124
of momentum, 135, 137
Critical angle, 1171–1172, 1181
Crossed wires, 985
Crystal lattice, 1214
Crystalline, 1214
Crystal monochromator, 1216
Current-carrying rods, 949
Current-carrying wire, 947–948, 977–978, 980, 983–985, 988–989, 993–996
Current density, 1093–1094
Current loops, spin magnetism and, 978–980
Current loop torque, 981–982
Currents
alternating, 1122–1124
magnetic fields due to, 993–996
magnetism and, 954–956
resistance and, 1082–1084
Curvature, radius of, 630–631
Curve up/curve down, 88
Cutoff angular frequency, 1144
Cycle, 745
Cycloid, 405
Cylindrical symmetry, 853–854

D

Damped oscillations, 527–529
Damping, 506
Damping coefficient, 528
Damping force, 527–528
DC circuits, 1136
De Broglie, Louis, 1218
De Broglie wavelength, 1218
Deceleration, 87
Decimal places, 37
Decomposition
of forces, 319–322
of vectors, 317–319
Deductive reasoning, 22
Deformable objects, 271
Deformation
coherent and incoherent, 216–218
elastic restoring force and, 520
force and, 251, 256
in fluids, 617
friction and, 323
of macro-sized objects, 271
pressure and, 640
in simple harmonic motion, 520, 568
of spring scale, 435
Degree of freedom, 710
Degrees Celsius, 708
Degrees Fahrenheit, 708
Demagnetized clip, 943
Density, 33
Depletion zone, 1130
Destruction

to change value of extensive quantity, 123–124
of momentum, 135, 137
Destructive interference, 549
Desynchronization of twin clocks, 472
Diagrams, electric field, 816–817
Dielectric constant, 928–931
Dielectrics
capacitor with, 930–931
in charge separation and storage, 915–917
Gauss's law in, 931–932
properties of, 928
Diffraction, 592–594
of light, 1208–1210
at single-slit barrier, 1229–1231
X-ray, 1213–1217
Diffraction gratings, 1210–1213, 1226–1227
Diffraction limit, 1233
Dilation, time, 479–484
Diodes, 1130–1132
Diopters, 1186
Dipole, 832–835
charge, 825–826
electrostatic potential and electric field due to, 899
field, 825–827
flux of, 851
induced, 834
moment, 821–822, 826–827, 833–835
oscillating, 1049–1053
permanent, 821, 822
right-hand dipole rule, 980
separation, 825–826
Dipole antenna, 1053
Direct current, 1122
Direct integration, 863
Direction
of current, 947, 1097
of time, 26
Discharged strip, 786
Dispersion, 1172
Displaced volume of fluid, 623, 639
Displacement, 55–57
calculus to determine, 106
distance traveled and, 57–61, 66, 68
magnitude of, 89–90
in motion in a plane
vectors in a plane, 316–319
work and friction, 324
work as product of two vectors, 331–335
of waves, 541–544, 546–550
in one dimension, 541–544, 546–550
in two and three dimensions, 581, 585–592
working with, 67–68
x component of, 56–57
Displacement current, in changing electric fields, 1055–1059
Displacement curve, 542
Displacement vector, 55–57, 66–70
Dissipative interactions, 216–219, 220, 221, 235–238
Distance, definition of, 66
Distance traveled, 57–61, 66, 68
Distributed forces, 300–303
Diverging lens, 1176–1177
Division, significant digits in, 37
Dopants, 1129
Doppler effect, 474, 602–607
Dozen, 33
Drag force, 528
Drain, 1133
Drift speed, 1093
Drift velocity, 1092–1093
Drude, P. K., 1092
Drude model, 1092

E

Earth
centripetal acceleration and, 426
gravitational force of, 426–430, 450
intensity of Sun's radiation on surface of, 1069–1070
North pole, pole attracted to, 942
orbit of, 426–427
satellites orbiting, 446–447
solar system data relative to, 430
surface of
gravitational potential energy near, 444–445
potential energy near, 232–234
weighing, 443
Earth-boulder system, 216
Earth reference frame, 177
center of mass and, 193–194
convertible kinetic energy and, 198–199
definition of, 177
Galilean relativity and, 190, 191
inertial reference frames and, 179–181
principle of relativity and, 182–184, 186
relativity of motion and, 177–179
zero-momentum reference frame and, 186–187
Eddy currents, 1019
Effective weight, 437
Einstein, Albert, 177, 462
Elastic collision, 150, 154, 160–162
Elastic limit, 255
Elastic potential energy, 216
Elastic range, 255
Elastic restoring force, 520
Electrical breakdown, 913
Electrical charge, 225, 785–788
Electrical charge symmetry, 24
Electrical conductors, 790
Electrical insulator, 789
Electrical measuring instruments, 1101
Electric circuits, 1079–1120
AC, 1124–1128, 1150–1152
basic circuit, 1080–1082
current and resistance, 1082–1084
electric fields in conductors, 1088–1090
energy conversion in, 1080–1084, 1098, 1106–1108
junctions and, 1084–1087
multiloop circuits, 1084–1087, 1101–1106
power in, 1106–1108, 1150–1152
RC series, 1141–1145, 1147–1148
resistance and Ohm's law, 1092–1096
RLC series, 1145–1146, 1148
single-loop circuits, 1096–1101
Electric constant, 860
Electric dipole. *See* Dipole
Electric field diagrams, 816–817
Electric field energy
charge separation and storage in, 925–928
emf and, 926
Electric field lines, 846–848
density of, 848–849
field line flux, 850–852
field strength from, 849
of infinite charged plate, 847
properties of, 849
Electric field model, 814–816
Electric fields, 813–844. *See also* Gauss's law
acceleration in, 820
accompanying changing magnetic field, 1015–1016, 1025–1027
calculating, using Gauss's law, 864
changing, 1042–1071
displacement current, 1055–1059
electromagnetic energy, 1067–1071
electromagnetic waves, 1043, 1062–1067
fields of moving charged particles, 1046–1049
magnetic fields accompanying, 1043–1046
Maxwell's equations, 1059–1062
oscillating dipoles and antennas, 1049–1053
in charged conducting objects, 854–856
in charged particle, 958–963, 1046–1049
in conductors, 1088–1090
of continuous charge distributions, 827–832
creating
by uniformly charged disk, 830–831
by uniformly charged sphere, 831–832
by uniformly charged thin ring, 830
by uniformly charged thin rod, 829
cylindrical Gaussian surface in, 859–860
definition of, 814
diagrams, 816–817
dipole field, 825–827
dipoles, 832–835
forces and, 820–822
of infinite charged conducting plate, 866
of infinite charged sheet, 865–866
of infinite charged thin rod, 865
line integrals of, 983–984
lines, 846–848
magnitude of, 848
model, 814–816
moving particles in, 958–963
obtaining, from electrostatic potential, 897–899
superposition of, 817–820
surrounding charged insulated wire, 932–933
due to two charged particles, 825
in uniformly charged sphere, 864
uniform *versus* nonuniform, 820, 821–822
Electric flux, 858–860, 861
Electric force, 797
Electric interactions, 783–812
charge polarization, 793–794
Coulomb's law, 796–800
electrical charge, 785–788
force exerted by distributions of charge carriers, 800–803
mobility of charge carriers, 788–793
static electricity, 784–785

Electricity
gravity *versus,* 798
magnetism unified with, 963–965
Electric motors, 982
Electric potential energy, 879–880
Electric trajectory, 803
Electrodes, 917–919
Electrolyte, 917–919
Electromagnetic energy
in changing electric fields, 1067–1071
density, instantaneous values of, 1069
Electromagnetic induction, 1014, 1016, 1021
Electromagnetic interactions, 224, 225, 942, 952
Electromagnetic wave pulses, 1048–1049
Electromagnetic waves, in changing electric fields, 1043, 1062–1067
Electromagnetism, 846. *See also* Gauss's law
Electromotive force. *See* Emf (electromotive force)
Electronics, 1121–1163
AC circuits, 1124–1128
power in, 1150–1152
alternating currents, 1122–1124
diodes, 1130–1132
logic gates, 1134
RC series circuit, 1141–1145, 1147–1148
reactance, 1136–1141
resonance, 1148–1149
RLC series circuit, 1145–1146, 1148
semiconductors, 1128–1130
transistors, 1132–1134
Electron microscope, 1217
Electrons, 24
magnetic fields generated by moving, 997
mass of, 961–962
Electroscope, 790
Electrostatic equilibrium, 802, 854, 855, 856
Electrostatic field, 880–882, 884, 894–895
Electrostatic force, 797
Electrostatic potential
of continuous charge distributions, 895–897
electric field due to dipole and, 899
obtaining electric field from, 897–899
of a uniformly charged disk, 896–897
of a uniformly charged thin rod, 896
Electrostatic potential difference, 879, 882, 889–895
atomic, 891–892
definition of, 889
potential and, 890
between two points in electric field, calculating, 892
Electrostatics, 797, 878–907
calculating, 886–889
electric potential energy, 879–880
electrostatic work, 880–882
equipotentials, 882–884
obtaining electric field from potentials, 897–899
potential difference, 879, 889–895
potentials of continuous charge distributions, 895–897
Electrostatic work, 880–882
Electroweak interactions, 226
Elementary magnets, 943
Ellipses, 448
Emf (electromotive force)
definition of, 918
electric field energy and, 926
induced, in changing magnetic fields, 1021–1025
Emitter, 1132
Energy, 150–175
chemical, 156, 219, 221
closed systems, 156–158
coherent (or mechanical), 217–220, 221, 237
collisions, classification of, 150–151
elastic, 150, 154, 160–162
inelastic, 150, 154, 163–164
conservation of, 164–166
in closed system, 164–166
law of, 150, 153
special relativity and, 492–496
consumption, 221
convertible kinetic, 198–201
converting (*See* Energy conversion)
definition of, 743
degradation of (*See* Energy degradation)
delivering, 569
in electrostatics (*See* Electrostatics)
entropy and
dependence of energy on, 683–687
energy transfers and, constraints on, 761–764
equipartition of energy, 669–670
equipartition of, 669–670
explosive separations, 166–168
gravitational potential, 443–446
incoherent, 217–220, 221, 236, 237
inertia and, 477
input-output diagrams, 743–744
in interactions
energy dissipation, 216–219
potential energy, 215–216, 232–234
source energy, 219–222
internal, 153–156, 217
kinetic (*See* Kinetic energy)
magnetic, in changing magnetic fields, 1029–1031
mass of, 476–477
matter and, 474–477
nuclear, 219
periodic motion and, 505–507
photon, 1234–1238
potential
electric (*See* Electric potential energy)
in interactions, 215–216, 232–234
near Earth's surface, 232–234
quality of, 746–750
of simple harmonic oscillator, 516–519
solar, 219
of sound, 596–597
of sound waves, 596–597
storage (*See* Charge separation and storage)
stored solar, 219
of system, 153
thermal, 156, 221
torque and, 409–411
transferred thermally (*See* Energy transferred thermally)
transport in waves, 567–569
Energy conversion, 743–746
in circuits, 1080–1084, 1098, 1106–1108
energy degradation and, 743–746
energy law and, 743, 745, 761, 762, 766
by heat engines and heat pumps, 750–756
mechanical to thermal energy, 744–745
principle of relativity and, 184–186
quality of energy and, 746–750
in steady devices, 745–746
Energy degradation, 742–782
Brayton cycle, 770–772
Carnot cycle, 768–770
converting energy, 743–746
of energy quality, 749–750
entropy constraints on energy transfers, 761–764
heat engines, 750–754, 764–768
heat pumps, 754–756
quality of energy, 746–750
thermodynamic cycles, 756–759
upgrading energy, 754
Energy diagrams, 286–288
Energy dissipation, 216–219
Energy input, 743
Energy input-output diagrams, 743
Energy law, 293, 743, 745, 761, 762, 766
Energy output, 743
Energy transfer, 157. *See also* Energy conversion*;* Energy transferred thermally
Energy transferred thermally, 701–741
change in energy and work, 719–721
entropy change in ideal gas processes, 727–731
entropy change in nonideal gas processes, 731–732
heat capacity, 709–713
isentropic processes, 723
isobaric processes, 723–724
isochoric processes, 721–723
isothermal processes, 724–727
PV diagrams and processes, 713–717
temperature measurement, 706–709
thermal interactions, 702–706
Entropy, 664–700
change in ideal gas processes, 727–731
change in nonideal gas processes, 731–732
definition of, 679
dependence of
on energy, 683–687
on volume, 678–683
energy transfers and, constraints on, 761–764
equipartition
of energy, 669–670
of space, 670–672
macrostates and, 673–676
of monatomic ideal gas, 690–691
states, 665–669
Entropy diagrams, 748–749
Entropy gradient, 747, 748
Entropy law, 679
Environment, systems and, 122
Equation of motion, 259–262
Equilibrium
electrostatic, 802, 854, 855, 856

mechanical, 402
positions, 511–512
thermal, 685–687, 690
translational, 251–252, 402
Equipartition
of energy, 669–670
of space, 670–672
Equipotential lines, 882–883
Equipotentials, 882–884
Equipotential surfaces, 883–884, 897
Equipotential volume, 884
Equivalence, principle of, 438–440
Escape speed, 451
Evaporator, 759
Examples and exercises, overview of, 30
Explosive separations, 166–168
Extended free-body diagrams, 393–395
Extensive quantities, 123–124
definition of, 123
processes that can change, 123–124
External force, 267–269, 283
External interactions, 132–133
External torque, 401

F

Fahrenheit scale, 708, 709
Faraday's law, 1014–1015
Far-sightedness, 1186
Fermat's principle, 1173, 1182
Field, in long-range interactions, 223
Field-effect transistor, 1133–134
Field forces, 250
Field line density, 848–849
Field line flux, 850–852
Filter, 1144–1145
Fluid flow, 624–628
Fluids, 616–663
Bernoulli's equation, 644–647
buoyancy and, 622–624
definition of, 617
displaced volume of, 623
flow of, 624–628
blood, 650
laminar, 626–627, 644–650
streamline, 626–627
turbulent, 625
forces in, 617–622
gravity and, 635–640
pressure and, 635–640
streamlines, 626–627
surface effects and, 628–633
surface tension and, 651–653
viscosity and, 647–651
working with pressure, 640–644
Focal plane, 1174
Focal point, 1173, 1178, 1185, 1187–1188
Focus, 1173, 1178, 1185, 1187–1188
Force, 247–281. *See also* Vector sum of forces
circular motion and, 356–358
contact, 250, 253
decomposing, 319–322
definition of, 247, 261
displacement, 283–284, 294
distributed, 300–303
electric field and, 820–822
equation of motion, 259–262
exerted by distributions of charge carriers, 800–803
external, 267–269, 283
field, 250
in fluids, 617–622
free-body diagrams, 252–254
of friction, 248, 322, 324
gravitational, exerted by a sphere, 451–453
gravitational field force, 253
of gravity, 262–263, 434–437
Hooke's law, 263–265
identifying, 250–251
impulse and, 265–267
interaction pair in, 249–250, 262
internal, 267–269, 270
linear restoring, 264
momentum and, 247–248
net, 248
Newton's laws of motion and, 261
point of application and
in force displacement, 283–284
induced emf and, 1021
torque and, 389, 390, 391, 393
principle of superposition in, 248
pull, 255
push, 255
reciprocity of, 248–250
restoring, 255
in simple harmonic motion, restoring, 511–514
spring constant and, 264
springs and, 255–256
superposition of, 260
support, 255
in systems of many interacting objects, 269–271
in systems of two interacting objects, 267–269
tensile, 256
tension, 256–257
translational equilibrium, 251–252
variable, 300–303
Force displacement, 283–284, 294
Force of friction, 248, 322, 324
Force of gravity, 434–437
Ford, Kevin, 629
Fourier analysis, 509–511
Fourier series, 509
Fourier's theorem, 509–511
Fourier synthesis, 511
Free-body diagrams, 252–254
Free electrons, 792
Free fall
acceleration of object in, 89–91
definition of, 89
Free-fall equations, 99–102
Free rotation, 392–393
Friction, 117
coefficients of, 336–340
collisions and, 117–122
definition of, 117
force of, 248, 322, 324, 338
kinetic, 323, 340
motion in a plane and, 322–324
static, 323, 324, 336 (*See also* Static friction)
velocity-*versus*-time graphs for, 117
work and, 323–324
Frictional force, 248
Frictional power, 303–304
Fringe order, 1210
Fully charged capacitor, 912
Fundamental interactions, 224–226
definition of, 224
electromagnetic, 224, 225
electroweak, 226
gravitational, 224–225
residual electromagnetic, 225, 226
residual strong, 226
strong, 224, 226
weak, 224, 225–226

G

Galilean relativity, 189–193
Galilean transformation equations, 190
Galileo Galilei, 89, 102, 177, 261, 438
Gate, 1133–1134
Gauge particles, 223–224
Gauge pressure, 640
Gauss's law, 845–877. *See also* Electric field lines
applying, 862–866
charged conducting objects, 854–856
closed surfaces, 849–851
definition of, 846
deriving, 860–861
in dielectrics, 931–932
electric flux, 858–860, 956
field line density, 848–849
for magnetism, 957, 983
symmetry and Gaussian surfaces, 851–854
Glossary terms, overview of, 30
Gluons, 226
Gram, 32
Graphs, 29
acceleration-*versus*-time, 131
calibrating, 55–56
momentum-*versus*-time, 136, 140, 186
for motion types, 98
position-*versus*-time, 88, 90, 104, 356
for representing waves, 540–543
slope of curve in, 60
velocity-*versus*-time
for acceleration, 104, 131
for classification of collisions, 150
for explosive separation, 167
for friction, 117
for inertia, 118, 119, 120, 121
for momentum, 136, 140
for motion in one dimension, 62
for principal of relativity, 182
x component in, 56–57
Gratings, diffraction, 1210–1213, 1226–1227
Gravitational constant, 442–443
Gravitational field, 814
magnitude of, 816
Gravitational field force, 253
Gravitational force
exerted by a sphere, 451–453
zero, 796
Gravitational interactions, 224–225
Gravitational potential energy, 216, 234, 290, 443–446

Graviton, 224
Gravity, 426–460
  acceleration due to, 88–90, 99–100, 102
  angular momentum and, 431–434
  celestial mechanics, 446–451
  electricity *versus,* 798
  fluids and, 635–640
  force of, 262–263
  gravitational constant, 442–443
  gravitational force exerted by a sphere, 451–453
  gravitational potential energy, 216, 234, 443–446
  gravitational pull, 428–430, 435–436, 438
  map of Earth's oceans, 90
  motion determined by, 91
  Newton's law of, 426–429
    Coulomb's law and, 797
    field model and, 814–815
    Kepler's laws of planetary motion and, 432–433
  principle of equivalence, 438–440
  surface tension and, 629–630
  universal, 426–431
  weight, 434–437
  work and, 295
Gross, 33
Grounding, 790

H

Hair cells, 585
Half-life, 471
Half-wave antenna, 1053
Hall effect, 962–963
Harmonic waves, 546
Heat, 703, 713
Heat capacity, 709–713
Heat capacity ratio, 772
Heat engines, 750–754
  performance of, 764–768
Heat pumps, 754–756
Helium density, 34
Henry (H), 1028
Hidden assumptions, 21
High-pass filter, 1144–1145
Hole, 1129
Hooke, Robert, 264
Hooke's law, 263–265, 548
Horizontal component of motion, 316
Huygens, Christiaan, 592
Huygens' principle, 592–593
Hydraulic systems, 642–643
Hydrostatic pressure, 638
Hydrostatics, 643
Hyperbolas, 448
Hypothesis, 20–22
  definition of, 20
  formulation and testing of, 21–22

I

Ideal ammeter, 1101
Ideal gas, 678
Ideal gas law, 689
Ideal gas processes
  entropy change in, 727–731
  isentropic, 723
  isobaric, 723–724
  isochoric, 721–723
  isothermal, 724–727
Ideal voltmeter, 1101
Image distance, 1183
Images, forming, 1173–1178
Impedance of the load, 1144
Impulse
  force and, 265–267
  momentum and, 137
  rotational, 402
Impulse equation, 266–267
Impurities, 1129
Incident pulse, 550
Inclined planes, acceleration on, 102–103
Incoherent deformation, 216–218
Incoherent energy, 217–220, 221, 236, 237
Incompressible liquids, 619
Induced current, 1012–1019, 1022–123, 1025, 1027
Induced dipole, 834
Induced emf, in changing magnetic fields, 1021–1025
Induced magnetic field, 1017–1018
Inductance, 1027–1029
  calculating, 1028
  definition of, 1027–1028
  of solenoid, 1029
Inductive reasoning, 22
Inductor, 1028
Inelastic collision, 150, 154, 163–164
Inertia, 117–122
  collisions and, 117–122
  definition of, 120
  determinants of, 121–122
  energy and, 477
  gravity and, 429
  inertial standard, 127–128
  momentum and, 128–130
  ratio of, 120
  reduced, 199
  rotational (*See* Rotational inertia)
  as scalar, 127
  symbol for, 127
  unit of, 127
  velocity and, 120, 121–122
Inertial reference frames, 179–181, 183–185, 201–202
Inertial standard, 127–128
Initial phase, 517
In phase waves, 562–563, 587
Input
  to change value of extensive quantity, 123–124
  in isolated system, 132
  of momentum, 137
Input-output diagrams, energy, 743–744
Instantaneous acceleration, 103–106, 318–319
Instantaneous speed and velocity, 73
Instantaneous time rate of change, 259
Instantaneous values of electromagnetic energy density, 1069
Instantaneous velocity, 73–74, 318–319
Instant as a relative term, 56
Integral, 828
Intensity, waves in two and three dimensions and, 596–599
Intensity of the electromagnetic wave, 1069
Interacting objects
  conservation in, 181–184
  force in, 267–271
Interaction field, 814
Interaction pair, 249–250, 262
  of waves, 549
Interaction range, 222–224
Interactions, 212–245
  accelerations and, 228
  action at a distance, 814, 820
  attractive, 212
  definition of, 132, 212
  dissipative, 216–219, 220, 221, 235–238
  effects of, 212–215
  energy dissipation, 216–219
  external *vs.* internal, 132–133
  forces in
    systems of many interacting objects, 269–271
    systems of two interacting objects, 267–269
  fundamental (*See* Fundamental interactions)
  gauge particles in, 223–224
  interaction range, 222–224
  in isolated systems, 132–133
  long-range, 222–223
  momentum in, 132–133
  nondissipative, 220, 221, 229–232
  physical contact in, 222
  potential energy in, 215–216, 232–234
  repulsive, 212
  short-range, 223
  source energy in, 219–222
  strength of, 222
  strong, 224
Interference, waves and, 549, 587–592
Interference fringes, 1210
Internal energy, 153–156, 217
Internal force, 267–269, 270
Internal interactions, 132–133
Internal resistance, 1100
Internal torque, 401
International System, 32–35
Interpolating data, 57
Invariant, 467
Inverted reflected pulse, 550
Ion, 785, 791
Irreversible interactions, 216–219, 220, 221, 235–238
Irreversible processes, 153
Isentropes, 715, 717
Isentropic processes, 717, 723
Isobaric processes, 717, 723–724
Isochoric processes, 717, 721–723
Isochronous, 507
Isolated capacitor, 912
Isolated systems, 130–135
  choosing, 133–134
  definition of, 130
  interactions in, 132–133
  momentum of, 130, 132–133
Isothermal processes, 717, 724–727
Isotherms, 714–715

J

Joule (J), 283
Joules per coulomb (J/C), 889
Junctions, electric circuits and, 1084–1087

K

Kelvin scale, 708, 709
Kepler, Johannes, 426, 428, 432–433
Kepler's laws of planetary motion, 433
Kilogram, 32, 127
Kinematics, 54. *See also* Motion
  definition of, 54
  graphs for motion types, 98
  rotational, 362–367
Kinetic energy, 151–153
  closed-system approach to determine, 296
  conservation of energy and, 164–166
  convertible, 198–201
  definition of, 151, 217
  in Earth reference frames, 198–199
  in elastic collisions, 161–162, 183
  in explosive separations, 166–167
  in inertial reference frames, 183–185, 201–202
  internal energy and, 153–156
  open-system approach to determine, 296
  relativity and, 198–201
  thermal energy converted to, 157
  translational, 198–200
  in zero momentum reference frames, 186
  zero/nonzero, 177
Kinetic friction, 323
  coefficient of, 340

L

Laminar flow, 626–627, 644–650
Laplace's law, 631, 653
Latent heat, 713
Law
  of angular momentum, 402
  definition of, 20
  of gravity, Newton's, 426–429
    Coulomb's law and, 797
    field model and, 814–815
    Kepler's laws of planetary motion and, 432–433
  of inertia, 261
  of motion, Newton's, 261
    first, 261
    second, 261
    third, 261, 262
LC circuit, 1122
Lead-acid cell, 918
Leading zeros, 36
Leads, 1090
Length
  contraction, 474, 484–488
  definition of, 24
  in Galilean transformation equations, 190
  measurements at high speeds, 486
  object's proper, 473
  relaxed, 255
  standard unit for, 24
Length contraction, 474, 484–488
Lens, 1173
  converging, 1173–1177, 1183–1186
  diverging, 1176
  magnification of, 1184–1190
  simplified ray diagrams for, 1175
  strength, 1186
  thin, 1183–1186
Lensmaker's formula, 1193–1195
Lenz's law, 1016–1019
Light
  absorbed, 1166–1169
  diffracted, 1208–1210
  dispersed, 1172
  reflected, 1166–1169
  refracted, 1169–1172
  speed of, 466–467, 469, 474, 483–484
  transmitted, 1166–1169
Light clock, 469
Like charges, 786
Limiting process, 76
Linear charge density, 828
Linear mass density, 553, 565–566, 568–570
Linear restoring force, 264, 512–513, 519
Line integrals, 983–984
Logic gates, 1134
Longitudinal waves, 543
Long-range interactions, 222–223
Loop rule, 1097–1098
Loops
  definition of, 1080
  in multiloop circuits, 1084–1087, 1101–1106
  in single-loop circuits, 1096–1101
Lorentz factor, 480
Lorentz transformation equations, 488
Loudness, 585

M

Mach number, 609
Macrostates, entropy and, 673–676
Magnetic constant, 987
Magnetic dipole moment, torque and, 980–983
Magnetic energy
  in changing magnetic fields, 1029–1031
  in square toroid, 1031
Magnetic field meter, 955–956
Magnetic fields, 944–946, 955
  Ampère's law for calculating, 987
  in capacitors, 1057–1058
  changing, 1011–1041
    electric field accompanying changing magnetic field, 1015–1016, 1025–1027
    Faraday's law, 1014–1015
    induced currents in, 1012–1019, 1022–123, 1025, 1027
    induced emf, 1021–1025
    inductance, 1027–1029
    Lenz's law, 1016–1019
    magnetic energy in, 1029–1031
    moving conductors in magnetic fields, 1012–1014
  of charged particles in motion, 976–1010
    Ampère's law, 987–990
    Ampèrian paths, 983–985
    current loops and spin magnetism, 978–980
    magnetic dipole moment and torque, 980–983
    magnetic field of moving charged particle, 996–999
    magnetic fields due to currents, 993–996
    solenoids, 990–991
    source of the magnetic field, 977–978
    toroids, 991–992
  of current-carrying wire, 977–978, 980, 983–985, 988–989, 993–996
  due to currents, 993–996
  generating
    by circular arc of current-carrying wire, 995
    by a large current-carrying sheet, 989
    by long straight current-carrying wire, 988, 994
    by moving electron, 997
  line integrals of, 983–984
  moving particles in, 958–963
Magnetic flux, 956–958
Magnetic interactions, 941–975
  magnetic fields, 944–946
  magnetic flux, 956–958
Magnetic materials, 943
Magnetic monopoles, 943
Magnetic poles, 942–943
Magnetism
  charge flow and, 946–949
  current and, 954–956
  definition of, 942
  electricity unified with, 963–965
  Gauss's law for, 957, 983
  Hall effect, 962–963
  as relativistic effect, 950
  relativity and, 949–952
  spin, current loops and, 978–980
Magnification, 1184–1190
Magnitude
  of acceleration, 87, 88
  of displacement, 89–90
  of drift velocity, 1093
  of electric fields, 848
  of electrostatic force, 797
  of gravitational field, 816
  sign for, 65
  of vector product, 412
  of velocity, 87, 88, 91
Manometer, 640
Many-particle systems, work on, 296–299
Mass
  center of, 193–197 (*See* Center of mass)
  definition of, 127
  densities, 34, 624
  of electrons, 961–962
  of energy, 476–477
  gravitational interactions and, 224
  gravity and, 429
  measuring, 566
  of the object, 428
  reduced, 199
  spring, 476–477
Mass density, 34, 624
Mass spectrometer, 960–961
Mathematics in physics, 35–38
Matter, 24–25
  classifying, 222
  clustering of, 25
  energy and, 474–477
  properties of, defined, 222
  relative size of, 25
  states of, 713 (*See also* States)

Matter waves, 1217–1218
Maxwell, James Clerk, 1056
Maxwell-Ampère law, 1056
Maxwell's equations, 1059–1062
Measuring instruments, electrical, 1101
Mechanical (or coherent) energy, 217–220, 221, 237
Mechanical equilibrium, 402
Mechanical waves, 540
Mechanics, celestial, 446–451
Melting, specific transformation energy for, 713, 714
Meniscus, 632–633
Metal-slab capacitor, 914–915
Meter, 24
Meters cubed per second (m^3/s), 645
Meters per second (m/s), 32
Michelson, A. A., 466
Mirrors, spherical, 1190–1193
Mobility of charge carriers, 788–793
Model
  definition of, 21
  in developing hypothesis, 21
  of real-world happening, 55
Moiré pattern, 587–588
Mole, 33
Momentum, 128–148
  angular (*See* Angular momentum)
  bounce and, 137
  change in, 129–130
  in collisions, 129–130
    conservation of momentum and, 135–140
    friction and, 117–122
    inertial standard and, 127–128
    in isolated systems, 133
    momentum changes and, 129–130
  conservation law and, 152, 201
  conservation of
    in collisions, 129–130
    in interactions, 135–140
    law of, 152, 201, 261
    special relativity and, 488–492
  creation of, 135, 137
  definition of, 128–130
  destruction of, 135, 137
  force and, 247–248
  impulse and, 137
  inertia and, 128–130
  input or output of, 132, 137
  in interactions, 132–133
  interactions in, 132–133
  of isolated systems, 130, 132–133
  photon energy, 1234–1238
  symbol for, 128
  of systems, 128–130
  time rate of change in, 247–248, 252, 259, 261, 269
  translational equilibrium and, 251–252
  in two dimensions, 330–331
  units of, 128
  as vector, 128
  velocity and, 128–129
  x component of, 128–129, 135, 138
Momentum law, 137–138
Momentum-*versus*-time graphs, 136, 140, 186
Monatomic ideal gas
  entropy of, 690–691
  properties of, 687–689
Monochromatic, 1216
Moon
  centripetal acceleration and, 427
  gravitational force of, 426–430, 450
  orbit of, 426–427
Morley, E. W., 466
Motion
  in a circle (*See* Circular motion)
  with constant acceleration, 96–99
  coordinates in, 316
  equation of, 259–262
  gravity for determining, 91
  horizontal component of, 316
  Kepler's laws of planetary, 433
  kinematics graphs for, 98
  Newton's laws of, 261
  in one dimension, 54–85
    average speed and average velocity, 59–62
    instantaneous velocity, 73–74
    motion at constant velocity, 71–73
    position and displacement vector, 55–57, 66–70
    from reality to model, 54–55
    representing motion, 57–59
    scalar and vector, 64–66
    velocity as a vector, 70–71
  orbital, 426
  phase of, 517
  in a plane, 315–350
    coefficients of friction, 336–340
    collisions and momentum in two dimensions, 330–331
    decomposition of forces, 319–322
    friction and, 322–324
    projectile motion in two dimensions, 328–330
    straight as a relative term, 315–316
    vector algebra, 326–328
    vectors in, 316–319
    work as product of two vectors, 331–335
  projectile, in two dimensions, 328–330
  at rest, 87, 92, 98
  rolling, 404–409
  rotational equation of, 400
  translational, 393, 395–396, 404–405
  types
    acceleration-*versus*-time, 131
    kinematics graphs for, 98
    momentum-*versus*-time, 136, 140, 186
    velocity *versus* time, 96, 98, 104
  vertical component of, 316
  of waves, 541
Motion, relativity of, 177–179
Motion diagrams, 92–94, 98
Moving conductors, in magnetic fields, 1012–1014
Multiloop circuits, 1084–1087, 1101–1106
Multimeter, 1101
Multiple-slit interference, 1222–1227
Multiplication
  significant digits in, 37
  of vectors, 70

N

Natural philosophy, 20
Nearsightedness, 1186
Negative charge, 787
Negative electrical charge, 225
Negative work, 284–286
Net force, 248
Neutral buoyancy, 640
Neutral equilibrium position, 511
Neutrons, 24
Newton (N), 259–260
Newton, Isaac
  law of gravity, 426–429
    Coulomb's law and, 797
    field model and, 814–815
    Kepler's laws of planetary motion and, 432–433
  laws of motion, 261
    first, 261
    second, 261
    third, 261, 262
  newton (N) in honor of, 259
  on time and change, 26
Node, 550
Nonconvertible kinetic energy, 198–199
Nondissipative processes, 220, 221, 229–232
Nonelectrostatic, 918
Noninertial reference frames, 181
Nonmechanical waves, 540
Nonparaxial rays, 1175
Nonpolar dielectric, 915
Nonrelativistic speeds, 480
Non-SI units, 34–35
Non-SI units of measure, 32, 34–35
Nonuniform electric field, 820, 821–822
Nonviscous fluids, 625–626, 644–647
Nonzero acceleration, 92, 131, 822
Normal components, 320–321
North pole, Earth's, 942
North pole of magnet, 942
Npn-type bipolar transistor, 1132–1133
N-type silicon, 1132–1133
N-type wells, 1133–1134
Nuclear energy, 219
Number density, 33

O

Object distance, 1183
Objects
  charged conducting, 854–856
  deformable, 271
  escape speed of, 451
  interacting
    conservation in, 181–184
    force in, 267–271
  mass of, 428
  in mechanical equilibrium, 402
  proper length of, 473
  shape factor of, 407
  in translational equilibrium, 402
  vector sum of forces exerted on, 393
Observer, 177
Ohmmeter, 1101
Ohm's law, 1092–1096
Ohm Ω, 1095
Open-system approach, 296

Opposite charges, 786
Optical instruments, 1186–1190
Orbital motion, 426
Orbits
  elliptical, 448
  escape speed and, 451
  hyperbolic, 448
  parabolic, 448
  planetary, 427–430
  satellite, 446–450
Order of magnitude, 24
Order-of-magnitude estimates, 42–45
Order-of-magnitude rounding, 24
Origin, 54
Ørsted, Hans Christian, 946
Oscillating dipoles, 1049–1053
Oscillation phases, 517–518
Oscillations, 505
  damped, 527–529
  from interplay between inertia and restoring force, 514
  vertical, 523
Oscillator, torsional, 524
Out of phase waves, 562–563
Output
  to change value of extensive quantity, 123–124
  in isolated system, 132
  of momentum, 137

P

Parabola, 448
Parallel-plate capacitor, 911–914, 922
Paraxial approximation, 1186
Paraxial rays, 1175
Parity symmetry, 24
Partial derivative, 570
Particles, se also wave and particle optics
  charged
    in electric fields, 824–825
    in magnetic fields, 976–1010
  definition of, 293
  gauge, in interactions, 223–224
  many-particle systems, work on, 296–299
  moving charged particles
    in electric fields, 958–963, 1046–1049
    in magnetic fields, 958–963, 976–1010
  single, work on, 293–296
  subatomic, 20, 24–25
  work equation for, 294–295
  work on many-particle system, 296–299
  work on single particle, 293–296
Pascal (Pa), 635–636
Pascal, Blaise, 635
Pascal's principle, 621, 630
Path independence
  of change in potential energy, 230–231
  of gravitational potential energy, 234
Peak value, 1136
Pendulum, 513
  dissipation, 669–670
  energy, 665
  period of, 514
  simple, 513, 526, 527
Perfect vacuum, 635
Perihelion, 429
Periodic motion, 505–538
  damped oscillations, 527–529
  definition of, 505
  energy and, 505–507
    of a simple harmonic oscillator, 516–519
  Fourier's theorem, 509–511
  simple harmonic motion, 507–509
    energy of a simple harmonic oscillator, 516–519
    restoring forces in, 511–514
    springs and, 520–524
  torques, restoring, 524–527
Periodic waves, 546–547
Period of pendulum, 514
Permanent dipoles, 821
Permanent electric dipole, 822
Permeability constant, 987
Phase constant, 1138–1139
Phase of the motion, 517
Phase transition, 713, 714
Photoelectric effect, 1238
Photolithography, 1233
Photons, 1218–1220
  in electromagnetic interactions, 225
  energy and momentum of, 1234–1238
Physical quantities, 24
  symbols for, 32
  units of, 32–35
Physics
  definition of, 20
  mathematics in, 29, 35–38
Pion, 224
Pitch, 585
Planar symmetry, 853, 854
Planar wavefront, 583
Planck's constant, 1218
Planetary motion, Kepler's laws of, 433
Planetary orbits, 427–430
Plus sign in front of square root, 326, 327
Pluto's orbit, 429–430, 433, 434
Pnp-type bipolar transistors, 1132–1133
Point of application
  in force displacement, 283–284
  induced emf and, 1021
  torque and, 389, 390, 391, 393
Poiseuille's law, 650
Polar coordinates, 326
Polar dielectric, 915
Polarization, 793–794
  of polar dielectric, 915–917
  of the wave, 1053
Polar vectors, 398
Position vector, 55–57, 66–70
Position-*versus*-time graphs, 88, 90, 104
Positive charge, 787
Positive electrical charge, 225
Positive work, 284–286
Positron, 792
Potential difference. *See* Electrostatic potential difference
Potential difference, cell, 918
Potential energy
  definition of, 215–216, 217
  elastic, 216
  electric (*See* Electric potential energy)
  gravitational, 216, 234, 290, 443–446
  in interactions, 215–216, 232–234
  near Earth's surface, 232–234
  path independence of change in, 230–231
  of springs, 301–302
Power
  definition of, 303
  in electric circuits, 1106–1108, 1150–1152
  frictional, 303–304
  work and, 303–305
Power factor, 1152
Poynting, J. H., 1069
Poynting vector, 1069–1071
Prefixes, for units of measure, 32
Pressure
  definition of, 619
  difference across the surface, 631
  fluids and, 619, 635–640
  in liquids at rest, 643
  water and air, 621–622
  working with, 640–644
Pressure gauge, 640
Pressure ratio, 772
Priestley, Joseph, 796
Principal maxima, 1212–1213, 1225–1227
Principal rays, 1174–1175, 1191
Principle of causality, 26
Principle of equivalence, 438–440
Principle of relativity, 181–186. *See also* Relativity
  consequence of, 185–186
  formula, 185
Principle of superposition, 248, 828
Problem-solving strategy, 38–42
  devise plan, 39
  evaluate result, 40, 41
  execute plan, 39–40
  getting started, 39
Procedures, overview of, 30
Processes
  adiabatic, 703, 717
  constrained quasistatic, 717
  definition of, 153, 704
  entropy change in
    ideal gas processes, 727–731
    nonideal gas processes, 731–732
  extensive quantities changed by, 123–124
  irreversible, 153
  isentropic, 717, 723
  isobaric, 717, 723–724
  isochoric, 717, 721–723
  isothermal, 717, 724–727
  limiting, 76
  PV diagrams and, 713–717
  quasistatic, 704–706, 715–717
  reversible, 153
Projectile
  acceleration due to gravity, 91
  definition of, 90
  range of, 330
  trajectory of, 90–92
Projectile motion, 90–92
  in two dimensions, 328–330
Proper time interval, 463
Proportionality constant, 1093
Protons, 24
P-type silicon, 1132–1133

Pull force, 255
Pulse, of waves, 550–554
Push force, 255
PV diagrams and processes, 713–717

Q

Quality factor, 529
Quality of energy, 746–750
Quanta, 667, 711
Quantitative description, 24
Quantum mechanics, 21, 667
Quarks, 226
Quasistatic processes, 704–706, 715–717
Questions and problems, overview of, 30

R

Radial coordinate, 326
Radius of curvature, 630–631, 1190
Rarefaction, 584
Ratio of inertia, 120
Ray diagram, 1167, 1171, 1173
  simplified, 1175, 1176, 1177
Rayleigh's criterion, 1232
Ray optics, 1164–1206
  absorbed light and, 1166–1169
  definition of, 1165
  dispersion and, 1172
  Fermat's principle and, 1173, 1182
  images and, forming, 1173–1178
  lensmaker's formula and, 1193–1195
  optical instruments and, 1186–1190
  rays and, 1165–1166
  reflected light and, 1166–1169
  refraction and, 1169–1172
  Snel's law and, 1180–1182
  spherical mirrors and, 1190–1193
  thin lenses and, 1183–1186
  transmitted light and, 1166–1169
Rays, 1165–1166
RC series circuit, 1141–1145, 1147–1148
Reactance, 1136–1141
Reasoning, 22
Reciprocity, of force, 248–250
Rectangular (Cartesian) coordinates, 317, 326, 898
Rectifier, 1131–1132
Recurrence time, 668–669
Reduced inertia, 199
Reference axis, 54–55
Reference clock, 464
Reference direction for the current, 1097
Reference frames. *See also* Earth reference frame
  conservation in, 177
    in inertial reference frames, 180–181
    in interacting objects, 181–184
    relativity and, 201–202
    in zero-momentum reference frames, 186
  definition of, 177
  inertial, 179–181
  noninertial, 181
  simultaneous events in, 465
  subscript operations for, 192
  zero-momentum, 186–187
Reflected light, 1166–1169
Reflected pulse, 550
Reflection
  angle of, 1166–1167, 1169, 1171, 1173
  total internal, 1171–1172, 1181
  of wave pulse, 551–554
Reflection symmetry, 22–23
Refraction, 1169–1172
  angle of, 1170–1172, 1180–1181
Relative speed, 150
Relativistic speeds, 480
Relativity, 177–210
  center of mass, 193–197
  conservation laws and, 201–202
  convertible kinetic energy, 198–201
  Einstein's theories of, 177
  Galilean relativity, 189–193
  inertial reference frames, 179–181
  magnetism and, 949–952
  of motion, 177–179
  principle of, 181–186
  zero-momentum reference frame, 186–187
Relativity of simultaneity, 468, 474
Relaxed length, 255
Representations, 27–29
  visual, 27–29
Repulsive forces, 628
Repulsive interactions, 212
Residual electromagnetic interactions, 225, 226
Residual strong interactions, 226
Resistance
  current and, 1082–1084
  internal, 1100
  Ohm's law and, 1092–1096
Resistance, of circuit element, 1084
Resistor, 1090
Resolution, 1226
Resolution, limits of, 1231–1234
Resonance, 1148–1149
Resonance curve, 1149
Resonant angular frequency, 1149
Restitution, coefficient of, 163–164
Restoring forces, 255, 264, 512–513
Reversible interactions, 220, 221, 229–232
Reversible processes, 153
Right-hand current rule, 947–948, 955, 980
Right-hand dipole rule, 980
Right-hand force rule, 948, 980
Right-hand rule, 396–397
RLC series circuit, 1145–1146, 1148
Rolling motion, 404–409
  acceleration in, 406–407
  angular momentum and, 367–372
  dynamics of, 406–407
  rotational inertia and, 400–404
  static friction and, 407
  velocity in, 404–406
  without slipping, 405–406
Root-mean-square values, 1069
Rotation
  about a fixed axis, 392, 404
  free, 392–393
  vectorial nature of, 395–398
Rotational equation of motion, 400
Rotational impulse, 402
Rotational impulse equation, 402
Rotational inertia, 358–360
  of extended objects, 372–377
  rolling motion and, 400–404
Rotational kinematics, 362–367
Rotational motion, 352. *See also* Circular motion
Rotational symmetry, 22–23
Rotational velocity vectors, 398

S

Satellite orbits, 446–450
Scalar, 64–66
Scalar field, 816
Scalar product, 333–334
Scales, time, 26–27
Scientific method, 20–21
  definition of, 20
  hypothesis in, 20–21
  significant digits and, 35–38
Second, 26
Secondary maxima, 1212–1213, 1225
Second derivative, 104
Second law of thermodynamics, 676
Self-quiz, overview of, 30
Semiconductors, 1128–1130
Shape factor of objects, 407
Shear stress, 617, 618–619, 625, 647
Shock waves, 607–609
Short circuit, 1086–1087
Short-range interactions, 223
Signal delay, 463
Significant digits, 35–38
Simple harmonic motion
  definition of, 505
  energy of simple harmonic oscillator, 516–519
  linear restoring force and, 508
  periodic motion and, 508
  restoring forces in, 511–514
  springs and, 520–524
Simple harmonic oscillator, 507, 516–519
Simple harmonic oscillator equation, 518–519
Simple pendulum, 513, 526, 527
Simultaneity detector, 465
Simultaneous as relative term, 465–469
Single-loop circuits, 1096–1101
Single particle, work on, 293–296
Single-slit barrier, diffraction at, 1229–1231
Single-variable integration, 859
Sinusoidal curves, 507
Sinusoidal electromagnetic waves, 1069–1070
Sinusoidal waves, 546, 547, 561, 571
SI units, 32–35. *See also* Units of measure
  for conductivity A/(V · m), 1093–1094
  coulomb (C), 797–798
  of damping coefficient (kg/s), 528
  henry (H), 1028
  joule (J), 283
  joules per coulomb (J/C), 889
  magnetic flux (T · m^2), 956
  meters cubed per second (m^3/s), 645
  newton (N), 259–260
  non-SI units accepted for use along with, 34–35
  pascal (Pa), 635–636
  prefixes, 32, 33
  for resistance (ohm Ω), 1095

seven base units, 32
in solar system data relative to Earth, 430
specific heat capacity (J/K · kg), 709
tesla N/(A · m), 955
volt (V), 889
watt (W), 303
Slipping, 323
rolling motion without, 405–406
Slope of the curve, 59
Small-angle approximation, 1186
Snel's law, 1180–1182
Snel van Royen, Willebrord, 1181
Solar energy, 219
Solar system data, 430
Solenoid, 990–991
inductance of, 1029
Sonic boom, 608–609
Sound waves
energy of, 596–597
perception of, 585
properties of, 585
in two and three dimensions, 584–587
Source energy, 219–222
Source of the field, 814
Source of the field, 814
South pole of magnet, 942
Space, equipartition of, 670–672
Space-time, 469–474
Space-time interval, 474, 483
Special relativity, 21, 462–503
conservation of energy and, 492–496
conservation of momentum and, 488–492
length contraction and, 484–488
matter and energy in, 474–477
simultaneous as relative term, 465–469
space-time and, 469–474
time dilation and, 479–484
time measurements and, 462–465
Specific heat capacity (J/K · kg), 709
Specific transformation energy, 713
for melting, 713, 714
for vaporization, 713, 714
Speed
average, 59–62, 71
circular motion at constant, 352–356
instantaneous, 73
unit of, 32
wave, 564–566
of waves, 541
Speed, average, 59–62
Speed of light, 466–467, 469, 474, 483–484
Sphere, gravitational force exerted by, 451–453
Spherical capacitor, 924
Spherical mirrors, 1190–1193
Spherical symmetry, 852, 853
Spherical waves, 582
Spin magnetism, current loops and, 978–980
Spots, 1216
Spring constant, 264
Spring mass, 476–477
Springs
compressing, 288, 301
elastic limit of, 255
elastic restoring force and, 520
force exerted by, 255–256
potential energy of, 301–302
simple harmonic motion and, 520–524
Spring scales, 435–437, 438
Square root, plus sign in front of, 326, 327
Square toroid, 992, 1031
Stable equilibrium position, 511
Standing waves, 562–564
States
basic, number of, 678
configuration of, 216
entropy and, 665–669
macrostates, 673–676
of matter, 617, 713
Static electricity, 784–785
Static friction
coefficient of, 336–340
definition of, 323
maximum force of, 336–338
rolling motion and, 407, 411
Steady device, 745
Stiffness, 255
Stopping potential difference, 1236
Stored solar energy, 219
Straight as a relative term, 315–316
Streamline flow, 626–627
Strong interactions, 224, 226
Subatomic particles, 20, 24–25
Subscript operations, 192
Subsonic waves, 585
Subtraction
significant digits in, 37
of vectors, 69–70, 317
Sun
centripetal acceleration and, 427–428
energy from, 219
intensity of radiation on surface of Earth, 1069–1070
planetary orbits around, 429–432, 433
Superposition, principle of, 248, 828
Superposition of electric field, 817–820
Superposition of forces, 260
Superposition of waves, 548–550
Support force, 255
Surface charge density, 828
Surface effects, fluids and, 628–633
Surface integral, 859
Surfaces
closed, 849–851
symmetry and Gaussian, 851–854
Surface tension of fluids, 651–653
in contact with air, 651
gravity and, 629–630
Surplus charge, 787–788, 855
Symbols
for acceleration, 87
for inertia, 127
for momentum, 128
for physical quantities, 32
summarized, 75
Symmetry, 22–24
circular, 977–978
cylindrical, 853–854
definition of, 22
electrical charge, 24
Gauss's law and, 853
parity, 24
planar, 853, 854
reflection, 22–23
reflective, 22–23
rotational, 22–23
spherical, 852, 853
symmetry operations to determine, 853
translational, 23
Synchronized clocks, 462, 464–465
System diagram, 123, 133–134
Système International (SI), 32–35. *See also* SI units
Systems, 122–125
boundaries of, 122
center of mass of, 193–197
choice of, 289–291
choosing, 122–123
configuration state of, 216
constraints on, 123–124
definition of, 122
energy of, 153
environment and, 122
extensive quantities in, 123–124
external *vs.* internal interactions and, 132–133
force in
systems of many interacting objects, 269–271
systems of two interacting objects, 267–269
impulse delivered to, 137
intensive quantities in, 123
isochronous, 507
isolated, 130–135 (*See also* Isolated systems)
momentum of, 128–130
nonzero acceleration, 131
work on many-particle, 296–299

T

Tangential components, 320–321
Temperature conversions, 709
Temperature measurement, 706–709
of absolute temperature, 708
scales, 708–709
Celsius scale, 708, 709
comparison of, 709
Fahrenheit scale, 708, 709
Kelvin scale, 708, 709
Tensile forces, 256
Tensile stress, 617, 618, 619
Tension, 256–257
Tesla N/(A · m), 955
Theory, definition of, 20
Theory of special relativity, Einstein's, 462
Thermal energy, 156, 221. *See also* Energy transferred thermally
Thermal equilibrium, 685–687, 690
Thermal interactions, 702–706
Thermally insulated gas, 703
Thermal reservoir, 725
Thermodynamic cycles, 756–759
Thermodynamics, 702
Thin-film interference, 1227–1229
Thin lenses, 1183–1186
Threshold of hearing, 596
Threshold of pain, 596
Timbre, 585

Time, 26–27
absolute, 26
change and, 26
definition of, 26
derivatives, 75
direction of, 26
intervals, 75, 190, 463
irreversible flow of, 26
measurements, 462–465, 486
recurrence, 668–669
scales, 26–27
space-time, 469–474
units of, 26
Time-dependent wave function, 556–557
Time dilation, 470, 479–484
Time-independent wave function, 556
Time rate of change, in momentum, 247–248, 252, 259, 261, 269
Tonal quality, 585
Toroid, 991–992
square, 992, 1031
Torque, 389–424
algebraic sign of, 400
angular momentum and, 389–392
conservation of, 400–404
definition of, 389
energy and, 409–411
extended free-body diagrams, 393–395
external, 401
free rotation, 392–393
internal, 401
magnetic dipole moment and, 980–983
point of application and, 389, 390, 391, 393
restoring, 524–527
rolling motion, 404–409
vectorial nature of rotation, 395–398
vector product, 411–413
Torsional constant, 524–525
Torsional oscillator, 524
Total internal reflection, 1171–1172, 1181
Totally inelastic collision, 150–155, 163, 164, 166
Trailing zeros, 38
Trajectory, 90–92
charged particle, 820–821
electric, 803
Transistors, 1132–1134
Translational equilibrium, 251–252, 402
Translational kinetic energy, 198–200
Translational motion, 393, 395–396, 404–405
Translational symmetry, 23
Translucent materials, 1166
Transmitted light, 1166–1169
Transmural pressure, 653
Transverse waves, 543
Turbulent flow, 625
Turning points, 508
Twin paradox, 472

U

Ultrasonic waves, 585
Uniform electric field, 820, 821–822
Unit conversions, 35
Units
of inertia, 127
of momentum, 128
Units of measure, 32–35
abbreviations for, 32
for acceleration, 87
non-SI, 32, 34–35
physical quantities for, 32
prefixes for, 32
SI, 32–35
significant digits and, 35–38
Unit vector, 65
Unit vector notation, 65
Universal gravity, 426–431
Universe
definition of, 24
matter and, 24–25
structure and size scales of, 24–25
time scales in, 26–27
Unpaired charge, 788
Unstable equilibrium position, 511
Upgrading energy, 754

V

Vacuum, 635
Vacuum gauge, 640
Valence electrons, 1128–1130
Van de Graaff, Robert J., 910
Van de Graaff energy, 927
Van de Graaff generators, 910–911, 913, 917, 927, 1122
Vaporization, specific transformation energy for, 713, 714
Variable force, 300–303
Vector algebra, 326–328
Vector bosons, 225
Vector field, 816
Vectorial nature of rotation, 395–398
Vector product, 411–413
Vectors
addition of, 69, 317
algebra, 326–328
axial, 398
bosons, 225
component, 326–327
decomposing, 317–319
displacement, 55–57, 66–70
momentum as, 128
multiplication of, 70
in a plane, 316–319
polar, 398
position, 55–57, 66–70
Poynting, 1069–1071
product, 411–413
rotational velocity, 398
rotation of, 395–398
scalar and, 64–66
scalar product of two, 333–334
subtraction of, 69–70, 317
unit, 65
velocity as, 70–71
work as product of two, 331–335
x component of, 65–66
zero, 67
Vector sum of forces
definition of, 248
in equation of motion, 259–261
exerted on an object, 393
free-body diagrams, 252–254, 255, 256
in impulse equation, 266–267
in Newton's law of motion, 261
in systems of many interacting objects, 269, 271
in systems of two interacting objects, 268
translational equilibrium and, 251–252
Vector sum of horizontal displacement, 316–317
Velocity
average, 59–62, 71
center-of-mass, 195
changes in
acceleration and, 87–88
linear, 91
x component of, 105
constant, 87, 88, 90, 96, 98, 179
inertia and, 120, 121–122
instantaneous, 73–74, 318–319
magnitude of, 87, 88, 91
momentum and, 128–129
motion at constant, 71–73
in rolling motion, 404–406
as a vector, 70–71
x component of, 119–120, 121, 128, 135, 138
of zero-momentum reference frame, 186
Velocity gradient, 648
Velocity selector, 961
Velocity-*versus*-time graphs, 131, 140
for acceleration, 104, 131
for classification of collisions, 150
for explosive separation, 167
for friction, 117
for inertia, 118, 119, 120, 121
for momentum, 136, 140
for motion in one dimension, 62
for principal of relativity, 182
VENUS mnemonic, 40
Vertical component of motion, 316
Vertical oscillations, 523
Vibration, 505
Viscosity of fluids, 618, 625, 647–651
Visual representations, 27–29
Volt (V), 889
Volta, Alessandro, 889
Voltic cells, charge separation and storage in, 917–919
Voltmeter, 889–890, 1101, 1151
Volume
entropy and, 678–683
order-of-magnitude estimate of, 42–44
Volume charge density, 828
Volume flow rate, 645
Vortices, 625

W

Watt (W), 303
Wave equation, 569–571
Wavefronts, 581–583
Wave function of the wave, 581
Wavelength, 547
Wavelet, 592
Wave number of motion, 560
Wave propagation, 544–548
medium for, 540
Wave pulse, 1048
Waves

in one dimension, 539–579
amplitude of, 559, 569
boundary effects, 550–554
definition of, 540
displacement of, 541–544, 546–550
energy transport in waves, 567–569
harmonic, 546
interference and, 549
longitudinal, 543
mechanical, 540, 543
motion of, 541
nonmechanical, 540
out of phase, 562–563
periodic, 546–547
in phase, 562–563, 587
pulse of, 550–554
representing waves graphically, 540–543
sinusoidal, 546, 547, 561, 571
speed of, 541
standing waves, 562–564
superposition of waves, 548–550
transverse, 543
traveling, 558–559
wave equation, 569–571
wave functions, 556–562
wave propagation, 540, 544–548
wave speed, 564–566
polarization of, 1053
in two and three dimensions, 580–615
amplitude of, 583, 589, 592
beat frequency, 601
beating, 599–601
coherent sources, 589–592
compressions and rarefactions in longitudinal waves, 586
difference between, 583
diffraction, 592–594
displacement of, 581, 585–592
Doppler effect, 602–607
energy of sound, 596–597
intensity, 596–599
interference, 587–592
shock waves, 607–609
sound, 584–587, 596–597
spherical waves, 582
wavefronts, 581–583
wave function of the wave, 581
Waves, particle optics and, 1207–1246
circular apertures and limits of resolution, 1231–1234
diffraction
gratings, 1210–1213, 1226–1227
of light, 1208–1210
at single-slit barrier, 1229–1231
X-ray, 1213–1217
matter waves, 1217–1218
multiple-slit interference, 1222–1227
photons, 1218–1220
energy and momentum of, 1234–1238
thin-film interference, 1227–1229
Weak charge, 225
Weak interactions, 224, 225–226
Weber, Wilhelm, 956
Weighing devices, 435
Weight
of Earth, 443
gravitational force and, 434–437
notation of, 437
Wheatstone bridge, 1105–1106
Work, 283–313
concept of, 283
electrostatic, 880–882
energy diagrams, 286–289
equation for particles, 294–295
force displacement, 283–284, 294
friction and, 323–324
gravity and, 295
on many-particle system, 296–299
positive and negative, 284–286
power, 303–305
as product of two vectors, 331–335
on single particle, 293–296
system choice, 289–291
thermal energy transfer and, 719–721
variable and distributed forces, 300–303
Work function, 1237

X

X axis, 54–55
X component
of acceleration
average acceleration and, 87, 96, 104
constant acceleration and, 96
curvature of *x(t)* curve related to, 88
due to gravity, 90, 100
inclined planes and, 102–103
instantaneous acceleration and, 104–105
in projectile motion, 92
algebraic sign of, 57, 65–66
definition of, 56
of displacement, 56–57
in graphs, 56–57
of momentum, 128–129, 135, 138
of position, 69, 75
symbols, 75
of vector, 65–66
of velocity, 119–120, 121, 128, 135, 138
X coordinates, 54–55, 316
X-ray crystallography, 1217
X-ray diffraction, 1216
X(t) curve, 59

Y

Y coordinate, 316

Z

Zero gravitational force, 796
Zero-momentum reference frames, 186–187
Zeros
leading, 36
trailing, 38
without units, 38
Zeroth-order bright fringe, 1210
Zero vector, 67

This page intentionally left blank

SI Units, Useful Data, and Unit Conversion Factors

The seven base SI units		
Unit	**Abbreviation**	**Physical quantity**
meter	m	length
kilogram	kg	mass
second	s	time
ampere	A	electric current
kelvin	K	thermodynamic temperature
mole	mol	amount of substance
candela	cd	luminous intensity

Some derived SI units			
Unit	**Abbreviation**	**Physical quantity**	**In terms of base units**
newton	N	force	$kg \cdot m/s^2$
joule	J	energy	$kg \cdot m^2/s^2$
watt	W	power	$kg \cdot m^2/s^3$
pascal	Pa	pressure	$kg/m \cdot s^2$
hertz	Hz	frequency	s^{-1}
coulomb	C	electric charge	$A \cdot s$
volt	V	electric potential	$kg \cdot m^2/(A \cdot s^3)$
ohm	Ω	electric resistance	$kg \cdot m^2/(A^2 \cdot s^3)$
farad	F	capacitance	$A^2 \cdot s^4/(kg \cdot m^2)$
tesla	T	magnetic field	$kg/(A \cdot s^2)$
weber	Wb	magnetic flux	$kg \cdot m^2/(A \cdot s^2)$
henry	H	inductance	$kg \cdot m^2/(A^2 \cdot s^2)$

SI Prefixes					
10^n	**Prefix**	**Abbreviation**	10^n	**Prefix**	**Abbreviation**
10^0	—	—			
10^3	kilo-	k	10^{-3}	milli-	m
10^6	mega-	M	10^{-6}	micro-	μ
10^9	giga-	G	10^{-9}	nano-	n
10^{12}	tera-	T	10^{-12}	pico-	p
10^{15}	peta-	P	10^{-15}	femto-	f
10^{18}	exa-	E	10^{-18}	atto-	a
10^{21}	zetta-	Z	10^{-21}	zepto-	z
10^{24}	yotta-	Y	10^{-24}	yocto-	y

Values of fundamental constants

Quantity	Symbol	Value
Speed of light in vacuum	c_0	3.00×10^{8} m/s
Gravitational constant	G	6.6738×10^{-11} N · m^2/kg^2
Avogadro's number	N_A	$6.02214076 \times 10^{23}$ mol^{-1}
Boltzmann's constant	k_B	1.381×10^{-23} J/K
Charge on electron	e	1.60×10^{-19} C
Permittivity constant	ϵ_0	$8.85418782 \times 10^{-12}$ C^2/(N · m^2)
Permeability constant	μ_0	$4\pi \times 10^{-7}$ T · m/A
Planck's constant	h	6.626×10^{-34} J · s
Electron mass	m_e	9.11×10^{-31} kg
Proton mass	m_p	1.6726×10^{-27} kg
Neutron mass	m_n	1.6749×10^{-27} kg
Atomic mass unit	amu	1.6605×10^{-27} kg

Other useful numbers

Number or quantity	Value
π	3.1415927
e	2.7182818
1 radian	57.2957795°
Absolute zero ($T = 0$)	−273.15°C
Average acceleration g due to gravity near Earth's surface	9.8 m/s^2
Speed of sound in air at 20 °C	343 m/s
Density of dry air at atmospheric pressure and 20 °C	1.29 kg/m^3
Earth's mass	5.97×10^{24} kg
Earth's radius (mean)	6.38×10^{6} m
Earth–Moon distance (mean)	3.84×10^{8} m

Unit conversion factors

Length

1 in. = 2.54 cm (defined)

1 cm = 0.3937 in.

1 ft = 30.48 cm

1 m = 39.37 in. = 3.281 ft

1 mi = 5280 ft = 1.609 km

1 km = 0.6214 mi

1 nautical mile (U.S.) = 1.151 mi = 6076 ft = 1.852 km

1 fermi = 1 femtometer (fm) = 10^{-15} m

1 angstrom (Å) = 10^{-10} m = 0.1 nm

1 light − year (ly) = 9.461×10^{15} m

1 parsec = 3.26 ly = 3.09×10^{16} m

Volume

1 liter (L) = 1000 mL = 1000 cm^3 = 1.0×10^{-3} m^3
=1.057 qt (U.S.) = 61.02 $in.^3$

1 gal (U.S.) = 4 qt (U.S.) = 231 $in.^3$ = 3.785 L = 0.8327 gal (British)

1 quart (U.S.) = 2 pints (U.S.) = 946 mL

1 pint (British) = 1.20 pints (U.S.) = 568 mL

1 m^3 = 35.31 ft^3

Speed

1 mi/h = 1.4667 ft/s = 1.6093 km/h = 0.4470 m/s

1 km/h = 0.2778 m/s = 0.6214 mi/h

1 ft/s = 0.3048 m/s = 0.6818 mi/h = 1.0973 km/h

1 m/s = 3.281 ft/s = 3.600 km/h = 2.237 mi/h

1 knot = 1.151 mi/h = 0.5144 m/s

Angle

1 radian (rad) = 57.30° = 57°18'

1° = 0.01745 rad

1 rev/min (rpm) = 0.1047 rad/s

Time

1 day = 8.640×10^4 s

1 year = 365.242 days = 3.156×10^7 s

Mass

1 atomic mass unit (u) = 1.6605×10^{-27} kg

1 kg = 0.06852 slug

1 metric ton = 1000 kg

1 long ton = 2240 lbs = 1016 kg

1 short ton = 2000 lbs = 909.1 kg

1 kg has a weight of 2.20 lb where g = 9.80 m/s^2

Force

1 lb = 4.44822 N

1 N = 10^5 dyne = 0.2248 lb

Energy and work

1 J = 10^7 ergs = 0.7376 ft · lb

1 ft · lb = 1.356 J = 1.29×10^{-3} Btu = 3.24×10^{-4} kcal

1 kcal = 4.19×10^3 J = 3.97 Btu

1 eV = 1.6022×10^{-19} J

1 kWh = 3.600×10^6 J = 860 kcal

1 Btu = 1.056×10^3 J

Power

1 W = 1 J/s = 0.7376 ft · lb/s = 3.41 Btu/h

1 hp = 550 ft · lb/s = 746 W

1 kWh/day = 41.667 W

Pressure

1 atm = 1.01325 bar = 1.01325×10^5 N/m^2 = 14.7 $lb/in.^2$ = 760 torr

1 $lb/in.^2$ = 6.895×10^3 N/m^2

1 Pa = 1 N/m^2 = 1.450×10^{-4} $lb/in.^2$

Periodic Table of the Elements

Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Period 1	1 H 1.008																	2 He 4.003
Period 2	3 Li 6.941	4 Be 9.012											5 B 10.811	6 C 12.011	7 N 14.007	8 O 15.999	9 F 18.998	10 Ne 20.180
Period 3	11 Na 22.990	12 Mg 24.305											13 Al 26.982	14 Si 28.086	15 P 30.974	16 S 32.065	17 Cl 35.453	18 Ar 39.948
Period 4	19 K 39.098	20 Ca 40.078	21 Sc 44.956	22 Ti 47.867	23 V 50.942	24 Cr 51.996	25 Mn 54.938	26 Fe 55.845	27 Co 58.933	28 Ni 58.693	29 Cu 63.546	30 Zn 65.409	31 Ga 69.723	32 Ge 72.64	33 As 74.922	34 Se 78.96	35 Br 79.904	36 Kr 83.798
Period 5	37 Rb 85.468	38 Sr 87.62	39 Y 88.906	40 Zr 91.224	41 Nb 92.906	42 Mo 95.94	43 Tc (98)	44 Ru 101.07	45 Rh 102.906	46 Pd 106.42	47 Ag 107.868	48 Cd 112.411	49 In 114.818	50 Sn 118.710	51 Sb 121.760	52 Te 127.60	53 I 126.904	54 Xe 131.293
Period 6	55 Cs 132.905	56 Ba 137.327	71 Lu 174.967	72 Hf 178.49	73 Ta 180.948	74 W 183.84	75 Re 186.207	76 Os 190.23	77 Ir 192.217	78 Pt 195.078	79 Au 196.967	80 Hg 200.59	81 Tl 204.383	82 Pb 207.2	83 Bi 208.980	84 Po (209)	85 At (210)	86 Rn (222)
Period 7	87 Fr (223)	88 Ra (226)	103 Lr (262)	104 Rf (261)	105 Db (262)	106 Sg (266)	107 Bh (264)	108 Hs (269)	109 Mt (268)	110 Ds (271)	111 Rg (272)	112 Cn (285)	113 Nh (284)	114 Fl (289)	115 Mc (289)	116 Lv (292)	117 Ts (294)	118 Og (294)

Number of protons → 29
Symbol for element → Cu
63.546 ← Average atomic mass in g/mol. For elements having no stable isotope, value in parentheses is approximate atomic mass of longest-lived isotope.

Lanthanoids	57 La 138.905	58 Ce 140.116	59 Pr 140.908	60 Nd 144.24	61 Pm (145)	62 Sm 150.36	63 Eu 151.964	64 Gd 157.25	65 Tb 158.925	66 Dy 162.500	67 Ho 164.930	68 Er 167.259	69 Tm 168.934	70 Yb 173.04
Actinoids	89 Ac (227)	90 Th (232)	91 Pa (231)	92 U (238)	93 Np (237)	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)